AMERICAN
JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

PUBLISHED UNDER THE JOINT AUSPICES OF

THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

Volume LXXIX, Number 3
JULY, 1957

THE JOHNS HOPKINS PRESS
BALTIMORE 18, MARYLAND
U. S. A.

CONTENTS

The AMERICAN JOURNAL OF MATHEMATICS appears four times yearly.

The subscription price of the JOURNAL is $8.50 in the U. S.; $8.75 in Canada; and $9.00 in other foreign countries. The price of single numbers is $2.50.

Manuscipts intended for publication in the JOURNAL should be sent to Professor AUREL WINTNER, The Johns Hopkins University, Baltimore 18, Md.

Subscriptions to the JOURNAL and all business communications should be sent to THE JOHNS HOPKINS PRESS, BALTIMORE 18, MARYLAND, U. S. A.

THE JOHNS HOPKINS PRESS supplies to the authors 100 free reprints of every article appearing in the AMERICAN JOURNAL OF MATHEMATICS. On the other hand, neither THE JOHNS HOPKINS PRESS nor the AMERICAN JOURNAL OF MATHEMATICS can accept orders for additional reprints. Authors interested in securing more than 100 reprints are advised to make arrangements directly with the printers, J. H. FURST Co., 20 HOPKINS PLACE, BALTIMORE 1, MARYLAND.

The typescripts submitted can be in English, French, German or Italian and should be prepared in accordance with the instructions listed on the inside back cover of this issue.

Entered as second-class matter at the Baltimore, Maryland, Postoffice, acceptance for mailing at special rate of postage provided for in Section 1103, Act of October 3, 1917, Authorized on July 3, 1918.

PRINTED IN THE UNITED STATES OF AMERICA
BY J. H. FURST COMPANY, BALTIMORE, MARYLAND

ON c. s. s. COMPLEXES.*

By Daniel M. Kan.

1. **Introduction.** It was indicated in [3] how the usual notions of homotopy theory may be defined for cubical complexes which satisfy a certain extension condition. In the same manner (see [9]) these notions may be defined for complete semi-simplicial (c. s. s.) complexes which satisfy the following c. s. s. version of the extension condition. The notation used will be that of [2] except that the face and degeneracy operators will be denoted by ϵ^i and η^j (instead of $\epsilon_n{}^i$ and $\eta_n{}^j$).

Definition (1.1). A c. s. s. complex K is said to satisfy the *extension condition* if for every pair of integers (k, n) with $0 \leqq k \leqq n$ and for every n $(n-1)$-simplices $\sigma_0, \cdots, \sigma_{k-1}, \sigma_{k+1}, \cdots, \sigma_n \in K$ such that $\sigma_i\epsilon^{j-1} = \sigma_j\epsilon^i$ for $i < j$ and $i \neq k \neq j$, there exists an *n-simplex* $\sigma \in K$ such that $\sigma\epsilon^i = \sigma_i$ for $i = 0, \cdots, \hat{k}, \cdots, n$.

Let $\mathcal{S}$ be the category of c. s. s. complexes and c. s. s. maps and let $\mathcal{S}_E$ be its full subcategory generated by the c. s. s. complexes which satisfy the extension condition.

Many interesting c. s. s. complexes do not satisfy the extension condition; for example the finite c. s. s. complexes (finite = with only a finite number of non-degenerate simplices). The definitions of some homotopy notions, such as the homology groups, apply to all c. s. s. complexes, but the definition of the homotopy groups of [9], for instance, cannot be carried over to c. s. s. complexes which are not in $\mathcal{S}_E$.

In order to extend the definitions of all homotopy notions defined on the category $\mathcal{S}_E$ to the whole category $\mathcal{S}$ one needs what will be called an H-pair, i. e., a pair (Q, q) consisting of

(i) a functor $Q\colon \mathcal{S} \to \mathcal{S}_E$,

(ii) a natural transformation $q\colon E \to Q$ (where $E\colon \mathcal{S} \to \mathcal{S}$ denotes the identity functor), satisfying the following conditions:

(a) The functor Q maps homotopic maps into homotopic maps.

(b) Let $K \in \mathcal{S}_E$, then the map $qK\colon K \to QK$ is a homotopy equivalence.

* Received September 20, 1956.

(c) Let $K \in \mathcal{S}$ and let $f\colon QK \to QK$ be a map such that commutativity holds in the diagram

$$\begin{tikzcd} & QK \arrow[dd, "f"] \\ K \arrow[ur, "qK"] \arrow[dr, "qK"'] & \\ & QK \end{tikzcd}$$

Then f is a homotopy equivalence.

In view of condition (a) every homotopy notion on the category $\mathcal{S}_E$ yields by composition with the functor Q a homotopy notion on the whole category $\mathcal{S}$. Condition (b) implies that on the category $\mathcal{S}_E$ the homotopy notions induced by the functor Q coincide with the original ones. Condition (c) essentially ensures the uniqueness of the homotopy notions induced by Q; if (R, r) is another H-pair, then Q and R induce the same homotopy notions. In particular QK and RK have the same homotopy type, even if K does not satisfy the extension condition.

An example of an H-pair is the following. Let $S|\ |\colon \mathcal{S} \to \mathcal{S}_E$ be the functor which assigns to a c. s. s. complex K the simplicial singular complex $S|K|$ of the geometrical realization $|K|$ of K and let $j\colon E \to |\ |$ be the natural transformation which assigns to a c. s. s. complex K the natural embedding $jK\colon K \to S|K|$. Then it is readily seen that the pair $(S|\ |, j)$ is an H-pair.

Although the existence of an H-pair is sufficient in order to do homotopy theory on the whole category $\mathcal{S}$, it is sometimes convenient to have an H-pair which (unlike the pair $(S|\ |, j)$) may be defined in terms of c. s. s. complexes and c. s. s. maps only. Such an H-pair $(\mathrm{Ex}^\infty, \mathrm{e}^\infty)$ will be defined in this paper. A useful property of the functor $\mathrm{Ex}^\infty\colon \mathcal{S} \to \mathcal{S}_E$ is that it preserves fibre maps.

The main tool used in the definition of the functor Ex^∞ is what we call the *extension* $\mathrm{Ex}\,K$ of a c. s. s. complex K, which is in a certain sense dual to the *subdivision* $\mathrm{Sd}\,K$ of K. More precisely: let K and L be c. s. s. complexes, then there exists (in a natural way) a one-to-one correspondence between the c. s. s. maps $\mathrm{Sd}\,K \to L$ and the c. s. s. maps $K \to \mathrm{Ex}\,L$. In the terminology of [6] this means that the functor Ex is a right adjoint of the functor Sd.

The simplicial approximation theorem may be generalized to c. s. s. complexes roughly as follows: let $K, L \in \mathcal{S}$, K finite, then every continuous map $f\colon |K| \to |L|$ is homotopic with the geometrical realization of a c. s. s.

map $g\colon \mathrm{Sd}^n K \to L$ for some n. Using the adjointness of the functors Sd and Ex a dual theorem may be obtained which involves a c. s. s. map $h\colon K \to \mathrm{Ex}^n L$ instead of $g\colon \mathrm{Sd}^n K \to L$. This dual theorem may be strengthened as follows: let $K \in \mathcal{S}$ and $L \in \mathcal{S}_E$, then every continuous map $f\colon |K| \to |L|$ is homotopic with the geometrical realization of a c. s. s. map $h\colon K \to L$. It is essentially because of this property that, as far as homotopy theory is concerned, the c. s. s. complexes which satisfy the extension condition "*behave like topological spaces.*"

The paper is divided into two chapters. In Chapter I the definitions and results are stated; most of the proofs are given in Chapter II.

The results of this paper were announced in [5].

Chapter I. Definitions and results.

2. The standard simplices and their subdivision. For each integer $n \geqq 0$ let $[n]$ denote the *ordered set* $(0, \cdots, n)$. By a map $\alpha\colon [m] \to [n]$ we mean a *monotone function,* i. e., a function such that $\alpha(i) \leqq \alpha(j)$ for $0 \leqq i \leqq j \leqq m$.

For each integer $n \geqq 0$ the *standard n-simplex* $\Delta[n]$ is the c. s. s. complex defined as follows. A q-simplex of $\Delta[n]$ is a map $\sigma\colon [q] \to [n]$. For each map $\beta\colon [p] \to [q]$ the p-simplex $\sigma\beta$ is defined as the composite map

$$[p] \xrightarrow{\beta} [q] \xrightarrow{\sigma} [n].$$

For each map $\alpha\colon [m] \to [n]$ let $\Delta\alpha\colon \Delta[m] \to \Delta[n]$ be the c. s. s. map which assigns to a q-simplex $\tau \in \Delta[m]$ the composite map

$$[q] \xrightarrow{\tau} [m] \xrightarrow{\alpha} [n].$$

The *subdivision* of $\Delta\ [n]$ is the c. s. s. complex $\Delta'[n]$ defined as follows. A q-simplex of $\Delta'[n]$ is a sequence $(\sigma_0, \cdots, \sigma_q)$ where the σ_i are *non-degenerate* simplices of $\Delta[n]$ (i. e., the map $\sigma_i\colon [\dim \sigma_i] \to [n]$ is a monomorphism) and σ_i *lies on* σ_{i+1} (i. e.. $\sigma_i = \sigma_{i+1}\alpha$ for some α) for all i. For each map $\beta\colon [p] \to [q]$ we have $(\sigma_0, \cdots, \sigma_q)\beta = (\sigma_{\beta(0)}, \cdots, \sigma_{\beta(p)})$.

The *subdivision* of $\Delta\alpha$ is the c. s. s. map $\Delta'\alpha\colon \Delta'[m] \to \Delta'[n]$ given by $\Delta'\alpha(\tau_0, \cdots, \tau_q) = (\sigma_0, \cdots, \sigma_q)$, where σ_i is the unique non-degenerate simplex of $\Delta[n]$ for which (see [2]) there exist an epimorphism $\gamma_i\colon [\dim \tau_i] \to [\dim \sigma_i]$ such that commutativity holds in the diagram

$$\begin{array}{ccc} [\dim \tau_i] & \xrightarrow{\tau_i} & [m] \\ \downarrow{\scriptstyle \gamma_i} & & \downarrow{\scriptstyle \alpha} \\ [\dim \sigma_i] & \xrightarrow{\sigma_i} & [n] \end{array} \tag{2.1}$$

For each integer $n \geqq 0$ let $\delta[n] : \Delta'[n] \to \Delta[n]$ be the c. s. s. map which assigns to a q-simplex $(\sigma_0, \cdots, \sigma_q) \in \Delta'[n]$ the q-simplex $\sigma \in \Delta[n]$, i. e., the map $\sigma : [q] \to [n]$, given by $\sigma(i) = \sigma_i(\dim \sigma_i)$, $0 \leqq i \leqq q$.

LEMMA (2.2). *For each map $\alpha : [m] \to [n]$ commutativity holds in the diagram*

$$\begin{array}{ccc} \Delta[m] & \xrightarrow{\Delta\alpha} & \Delta[n] \\ \uparrow{\scriptstyle \delta[m]} & & \uparrow{\scriptstyle \delta[n]} \\ \Delta'[m] & \xrightarrow{\Delta'\alpha} & \Delta'[n] \end{array} \tag{2.2a}$$

Proof. It follows from the definitions that for every q-simplex $(\tau_0, \cdots, \tau_q) \in \Delta'[m]$ and each integer i with $0 \leqq i \leqq q$,

$$(\Delta\alpha \circ \delta[m])(\tau_0, \cdots, \tau_q)(i) = \alpha\tau_i(\dim \tau_i),$$

$$(\delta[n] \circ \Delta'\alpha)(\tau_0, \cdots, \tau_q)(i) = \delta[n](\sigma_0, \cdots, \sigma_q)(i) = \sigma_i(\dim \sigma_i),$$

where σ_i is the unique non-degenerate simplex of $\Delta[n]$ for which there exists an epimorphism γ_i such that commutativity holds in diagram (2.1). Because γ_i is onto,

$$\alpha\tau_i(\dim \tau_i) = \sigma_i\gamma_i(\dim \tau_i) = \sigma_i(\dim \sigma_i).$$

Hence commutativity holds in diagram (2.2a).

3. The extension of a c. s. s. complex. The *extension* of a c. s. s. complex K is the c. s. s. complex $\operatorname{Ex} K$ defined as follows. An n-simplex of $\operatorname{Ex} K$ is a c. s. s. map $\sigma : \Delta'[n] \to K$. For each map $\alpha : [m] \to [n]$ the m-simplex $\sigma\alpha$ is the composite map

$$\Delta'[m] \xrightarrow{\Delta'\alpha} \Delta'[n] \xrightarrow{\sigma} K.$$

Similarly the *extension* of a c. s. s. map $f : K \to L$ is the c. s. s. map $\operatorname{Ex} f$: $\operatorname{Ex} K \to \operatorname{Ex} L$ which assigns to every n-simplex $\sigma \in \operatorname{Ex} K$ the composite map

$$\Delta'[n] \xrightarrow{\sigma} K \xrightarrow{f} L.$$

Clearly the function Ex so defined is a covariant functor $\mathrm{Ex}\colon \mathcal{S} \to \mathcal{S}$. By Ex^n we shall mean the functor Ex applied n times.

For c. s. s. complex K define a monomorphism $eK\colon K \to \mathrm{Ex}\,K$ as follows. For every n-simplex $\sigma \in K$, $(eK)\sigma$ is the composite map

$$\Delta'[n] \xrightarrow{\delta[n]} \Delta[n] \xrightarrow{\phi\sigma} K,$$

where $\phi\sigma\colon \Delta[n] \to K$ is the unique map such that $\phi\sigma\alpha = \sigma\alpha$ for all $\alpha \in \Delta[n]$. It follows from Lemma (2.2) that the function e is a natural transformation $e\colon E \to \mathrm{Ex}$ (where $E\colon \mathcal{S} \to \mathcal{S}$ denotes the identity functor), i. e., for every c. s. s. map $f\colon K \to L$ commutativity holds in the diagram

$$\begin{array}{ccc} K & \xrightarrow{f} & L \\ \downarrow{\scriptstyle eK} & & \downarrow{\scriptstyle eL} \\ \mathrm{Ex}\,K & \xrightarrow{\mathrm{Ex}\,f} & \mathrm{Ex}\,L \end{array}$$

We shall denote by $e^nK\colon K \to \mathrm{Ex}^n K$ the composite monomorphism

$$K \xrightarrow{eK} \mathrm{Ex}\,K \xrightarrow{e(\mathrm{Ex}\,K)} \cdots \xrightarrow{e(\mathrm{Ex}^{n-1} K)} \mathrm{Ex}^n K.$$

LEMMA (3.1). *The functor* $\mathrm{Ex}\colon \mathcal{S} \to \mathcal{S}$ *maps homotopic maps into homotopic maps.*

The proof will be given in Section 9.

An important property of the functor Ex is that if it is twice applied to a c. s. s. complex K, then the resulting complex $\mathrm{Ex}^2 K$ partially satisfies the extension condition; if $\rho_0, \cdots, \rho_{k-1}, \rho_{k+1}, \cdots, \rho_n \in \mathrm{Ex}^2 K$ are n $(n-1)$-simplices which "match" and which are in the image of $\mathrm{Ex}\,K$ under the map $e(\mathrm{Ex}\,K)\colon \mathrm{Ex}\,K \to \mathrm{Ex}^2 K$, then there exists an n-simplex $\rho \in \mathrm{Ex}^2 K$ (not necessarily in the image of $\mathrm{Ex}\,K$) such that $\rho\epsilon^i = \rho_i$ for $i \neq k$. An exact formulation is given in the following lemma.

LEMMA (3.2). *Let* $K \in \mathcal{S}$. *Then for every pair of integers* (k, n) *with* $0 \leqq k \leqq n$ *and for* n $(n-1)$*-simplices* $\tau_0, \cdots, \tau_{k-1}, \tau_{k+1}, \cdots, \tau_n \in \mathrm{Ex}\,K$ *such that* $\tau_i\epsilon^{j-1} = \tau_j\epsilon^i$ *for* $i < j$ *and* $i \neq k \neq j$, *there exists an* n*-simplex* $\rho \in \mathrm{Ex}^2 K$ *such that* $\rho\epsilon^i = (e(\mathrm{Ex}\,K))\tau_i$ *for* $i = 0, \cdots, \hat{k}, \cdots, n$.

The proof will be given in Section 10.

Another useful property of the functor Ex is that it preserves fibre maps. This is stated in Lemma (3.4).

Definition (3.3). A c.s.s. map $f\colon K \to L$ is called a *fibre map* if for each pair of integers (k,n) with $0 \leqq k \leqq n$, for every n $(n-1)$-simplices $\tau_0, \cdots, \tau_{k-1}, \tau_{k+1}, \cdots, \tau_n \in K$ such that $\tau_i \epsilon^{j-1} = \tau_j \epsilon^i$ for $i < j$ and $i \neq k \neq j$ and for every n-simplex $\rho \in L$ such that $\rho \epsilon^i = f\tau_i$ for $i = 0, \cdots, \hat{k}, \cdots, n$, there exists an n-simplex $\tau \in K$ such that $f\tau = \rho$ and $\tau\epsilon^i = \tau_i$ for $i = 0, \cdots, \hat{k}, \cdots, n$. Let $\phi \in L$ be a 0-simplex. Then the counter image of ϕ and its degeneracies is called *the fibre of f over ϕ*. It is denoted by $F(f,\phi)$.

LEMMA (3.4). *Let $f\colon K \to L$ be a fibre map and let $\phi \in L$ be a 0-simplex. Then* $\operatorname{Ex} f\colon \operatorname{Ex} K \to \operatorname{Ex} L$ *is a fibre map and* $\operatorname{Ex}(F(f,\phi)) = F(\operatorname{Ex} f, (eL)\phi)$.

The proof will be given in Section 11.

Let $f\colon K \to \Delta[0]$ be a fibre map, then it follows readily from the fact that $\Delta[0]$ has only one simplex in every dimension that $K \in \mathcal{S}_E$. Conversely $K \in \mathcal{S}_E$ implies that the (unique) map $f\colon K \to \Delta[0]$ is a fibre map. As $\operatorname{Ex} \Delta[0] \approx \Delta[0]$ Lemma (3.4) thus implies

COROLLARY (3.5). *If $K \in \mathcal{S}_E$, then* $\operatorname{Ex} K \in \mathcal{S}_E$.

The following lemmas relate the homology groups of K and $\operatorname{Ex} K$ and, if $K \in \mathcal{S}_E$, their homotopy types.

LEMMA (3.6). *Let $K \in \mathcal{S}$. Then the map $eK\colon K \to \operatorname{Ex} K$ induces isomorphisms of the homology groups, i.e.,* $(eK)_*\colon H_*(K) \approx H_*(\operatorname{Ex} K)$.

The proof will be given in Section 12.

LEMMA (3.7). *Let $K \in \mathcal{S}_E$. Then the map $eK\colon K \to \operatorname{Ex} K$ is a homotopy equivalence.*

The proof will be given in Section 13.

4. The functor $\mathbf{Ex}^\infty$. Let K be a c.s.s. complex. Consider the sequence

$$K \xrightarrow{eK} \operatorname{Ex} K \xrightarrow{e(\operatorname{Ex} K)} \operatorname{Ex}^2 K \xrightarrow{e(\operatorname{Ex}^2 K)} \operatorname{Ex}^3 K \to \cdots$$

and let $\operatorname{Ex}^\infty K$ be the direct limit of this sequence. The n-simplices of $\operatorname{Ex}^\infty K$ then are the pairs (σ, q) where $\sigma \in \operatorname{Ex}^q K$ is an n-simplex; two n-simplices (σ, q) and $(\tau, p+q)$ are considered equal if and only if $(e^p(\operatorname{Ex}^q K))\sigma = \tau$. For each map $\alpha\colon [m] \to [n]$, $(\sigma,q)\alpha = (\sigma\alpha, q)$. Similarly for a c.s.s. map $f\colon K \to L$ let $\operatorname{Ex}^\infty f\colon \operatorname{Ex}^\infty K \to \operatorname{Ex}^\infty L$ be the induced map given by $f(\sigma,q) = (f\sigma, q)$. Clearly the function Ex^∞ so defined is a covariant functor.

For a c. s. s. complex K denote by $e^\infty K: K \to \mathrm{Ex}^\infty K$ the limit monomorphism

$$K \xrightarrow{eK} \mathrm{Ex}\, K \xrightarrow{e(\mathrm{Ex}\, K)} \cdots \to \mathrm{Ex}^\infty K$$

i. e., $(e^\infty K)\sigma = ((eK)\sigma, 1)$ for every simplex $\sigma \in K$. Naturality of the function e^∞ follows immediately from the naturality of e.

THEOREM (4.1). *The functor* Ex^∞ *maps homotopic maps into homotopic maps.*

The proof is similar to that of Lemma (3.1) (see Section 9), using Ex^∞ and e^∞ instead of Ex and e.

An important property of the functor Ex^∞ is:

THEOREM (4.2). $\mathrm{Ex}^\infty K \in \mathcal{S}_E$ *for all objects* $K \in \mathcal{S}$, i. e., Ex^∞ *is a functor* $E^\infty: \mathcal{S} \to \mathcal{S}_E$.

This follows immediately from Lemma (3.2) and the definition of Ex^∞.

Another useful property of the functor Ex^∞ is that it preserves fibre maps.

THEOREM (4.3). *Let* $f: K \to L$ *be a fibre map and let* $\phi \in L$ *be a 0-simplex. Then* $\mathrm{Ex}^\infty f: \mathrm{Ex}^\infty K \to \mathrm{Ex}^\infty L$ *is a fibre map and* $\mathrm{Ex}^\infty(F(f, \phi)) = F(\mathrm{Ex}^\infty f, (e^\infty L)\phi)$.

This follows immediately from Lemma (3.4).

We shall now relate the homology groups of K and $\mathrm{Ex}^\infty K$ and, if $K \in S_E$, their homotopy types.

THEOREM (4.4). *Let* $K \in \mathcal{S}$. *Then the map* $e^\infty K: K \to \mathrm{Ex}^\infty K$ *induces isomorphisms of the homology groups, i. e.,* $(e^\infty K)_*: H_*(K) \approx H_*(\mathrm{Ex}^\infty K)$.

This follows immediately from Lemma (3.6).

Similarly, Lemma (3.7) implies.

THEOREM (4.5). *Let* $K \in \mathcal{S}_E$. *Then the map* $eK: K \to \mathrm{Ex}^\infty K$ *is a homotopy equivalence.*

Let K be a c. s. s. complex which does *not* satisfy the extension condition. Then the homotopy type of $\mathrm{Ex}^\infty K$ cannot be related to the homotopy type of K because the latter has (not yet) been defined. However the homotopy type of $\mathrm{Ex}^\infty K$ may be related to K as follows:

THEOREM (4.6). *Let $K \in \mathcal{S}$ and let $f\colon \mathrm{Ex}^\infty K \to \mathrm{Ex}^\infty K$ be a c.s.s. map such that commutativity holds in the diagram*

$$\begin{tikzcd} & \mathrm{Ex}^\infty K \arrow[dd, "f"] \\ K \arrow[ur, "e^\infty K"] \arrow[dr, "e^\infty K"'] & \\ & \mathrm{Ex}^\infty K \end{tikzcd}$$

Then f is a homotopy equivalence.

The proof will be given in Section 14.

5. Homotopy notions induced on $\mathcal{S}$.

Definition (5.1). A pair (Q, q) where $Q\colon \mathcal{S} \to \mathcal{S}_E$ is a covariant functor and $q\colon E \to Q$ a natural transformation (E denotes the identity functor $E\colon \mathcal{S} \to \mathcal{S}$), is called an *H-pair* if the following conditions are satisfied.

(a) The functor $Q\colon \mathcal{S} \to \mathcal{S}_E$ maps homotopic maps into homotopic maps

(b) Let $K \in \mathcal{S}_E$. Then the map $qK\colon K \to QK$ is a homotopy equivalence

(c) Let $K \in \mathcal{S}$ and let $f\colon QK \to QK$ be a c.s.s. map such that commutativity holds in the diagram

$$\begin{tikzcd} & QK \arrow[dd, "f"] \\ K \arrow[ur, "qK"] \arrow[dr, "qK"'] & \\ & QK \end{tikzcd}$$

Then f is a homotopy equivalence.

Example (5.2). The pair $(\mathrm{Ex}^\infty, e^\infty)$ is an H-pair; this follows directly from Theorems (4.1), (4.5) and (4.6).

A more exact formulation of the statements about H-pairs made in the introduction will be given in Theorems (5.4), (5.5) and (5.8).

Definition (5.3). By a *homotopy notion* on the category $\mathcal{S}$ (resp. $\mathcal{S}_E$) with values in a category $\mathcal{G}$ we mean a functor $N\colon \mathcal{S} \to \mathcal{G}$ (resp. $N\colon \mathcal{S}_E \to \mathcal{G}$) such that for two maps $f, g \in \mathcal{S}$ (resp. $\mathcal{S}_E$) $f \simeq g$ implies $Nf = Ng$.

THEOREM (5.4). *Let $N\colon \mathcal{S}_E \to \mathcal{Z}$ be a homotopy notion on $\mathcal{S}_E$ and let (Q, q) be an H-pair. Then the composite functor*

$$\mathcal{S} \xrightarrow{\ Q\ } \mathcal{S}_E \xrightarrow{\ N\ } \mathcal{Z}$$

is a homotopy notion on $\mathcal{S}$.

This is an immediate consequence of condition (5.1a).

Let $J\colon \mathcal{S}_E \to \mathcal{S}$ be the inclusion functor and let $N\colon \mathcal{S}_E \to \mathcal{Z}$ be a homotopy notion on $\mathcal{S}_E$. We then want to compare the composite functor

$$\mathcal{S}_E \xrightarrow{\ J\ } \mathcal{S} \xrightarrow{\ Q\ } \mathcal{S}_E \xrightarrow{\ N\ } \mathcal{Z}$$

i.e., the restriction to $\mathcal{S}_E$ of the homotopy notion on $\mathcal{S}$ induced by the functor Q, with the original homotopy notion N on $\mathcal{S}_E$. The following theorem then asserts that these functors differ only by a natural equivalence.

THEOREM (5.5). *Let $N\colon \mathcal{S}_E \to \mathcal{Z}$ be a homotopy notion on $\mathcal{S}_E$ and let (Q, q) be an H-pair. Then the function $Nq\colon N \to NQJ$ is a natural equivalence.*

This follows immediately from condition (5.1b).

In order to prove the uniqueness of the homotopy notions on $\mathcal{S}$ induced by an H-pair (Q, q) we need the following lemma

LEMMA (5.6). *Let (Q, q) and (R, r) be H-pairs and let $K \in \mathcal{S}$. Then the maps $QrK\colon QK \to QRK$ and $RqK\colon RK \to RQK$ are homotopy equivalences.*

The proof will be given in Section 15; use will be made of condition (5.1c).

Let (Q, q) and (R, r) be H-pairs and consider the following commutative diagram

(5.7)
$$\begin{tikzcd}
QK \arrow[r, "qQK"] \arrow[d, "rQK"] & QQK \arrow[d, "QrQK"] & \\
RQK \arrow[r, "qRQK"] & QRQK & QK \arrow[ul, "QqK"'] \arrow[dl, "QrK"] \\
RK \arrow[r, "qRK"] \arrow[u, "RqK"] & QRK \arrow[u, "QRqK"] &
\end{tikzcd}$$

It follows from Lemma (5.6) and condition (5.1b) that all maps involved in diagram (5.7) are homotopy equivalences; application of a homotopy notion $N\colon \mathcal{S}_E \to \mathfrak{Z}$ to this diagram thus yields a diagram in $\mathfrak{Z}$ consisting only of equivalences. If we put $Q = R$ and $q = r$ then it follows from the commutativity of diagram (5.7) that

$$(NQqK)^{-1} \circ NqQK = (NqQK)^{-1} \circ NQqK \circ (NQqK)^{-1} \circ NqQK = i_{NQK}.$$

Consequently

$$\begin{aligned}(NRqK)^{-1} \circ NrQK &= (NqRK)^{-1} \circ NQrK \circ (NQqK)^{-1} \circ NqQK \\ &= (NqRK)^{-1} \circ NQrK.\end{aligned}$$

Hence the following uniqueness theorem holds.

THEOREM (5.8). *Let $N\colon \mathcal{S}_E \to \mathfrak{Z}$ be a homotopy notion on $\mathcal{S}_E$ and let (Q,q) and (R,r) be H-pairs. Then the function $h\colon NQ \to NR$ given by*

$$hK = (NRqK)^{-1} \circ NrQK = (NqRK)^{-1} \circ NQrK$$

is a natural equivalence.

6. The simplicial singular complex of the geometrical realization. We shall now use the results of Section 5 in order to compare the simplicial singular complex of the geometrical realization of a c.s.s. complex K with $\mathrm{Ex}^\infty K$.

Let $\mathfrak{A}$ be the category of topological spaces and continuous maps and let $|\ |\colon \mathcal{S} \to \mathfrak{A}$ be the *geometrical realization functor* which assigns to a c.s.s. complex K its geometrical realization $|K|$ in the sense of J. Milnor (see [8]); $|K|$ is a CW-complex of which the n-cells are in one-to-one correspondence with the non-degenerate n-simplices of K.

Let $S\colon \mathfrak{A} \to \mathcal{S}_E$ be the *simplicial singular functor* which assigns to a topological space X its simplicial singular complex SX (see [2]); an n-simplex of SX is any continuous map $\sigma\colon |\Delta[n]| \to X$ and for every map $\alpha\colon [m] \to [n]$ the n-simplex $\sigma\alpha$ is the composite map

$$|\Delta[m]| \xrightarrow{|\Delta\alpha|} |\Delta[n]| \xrightarrow{\sigma} X.$$

The functor S maps homotopic maps into homotopic maps.

For every c.s.s. complex K let $jK\colon K \to S|K|$ be the natural monomorphism which assigns to an n-simplex $\sigma \in K$ the simplex $|\phi\sigma|\colon |\Delta[n]|$

$\to |K|$ of $S|K|$, where $\phi_\sigma\colon \Delta[n] \to K$ is the unique c. s. s. map such that $\phi_\sigma \alpha = \sigma\alpha$ for all $\alpha \in \Delta[n]$.

The following results are due to J. Milnor ([8]).

THEOREM (6.1). *The functor* $|\ |\colon \mathcal{S} \to \mathcal{A}$ *maps homotopic maps into homotopic maps.*

COROLLARY (6.2). *The functor* $S|\ |\colon \mathcal{S} \to \mathcal{S}_E$ *maps homotopic maps into homotopic maps.*

THEOREM (6.3). *Let* $K \in \mathcal{S}_E$. *Then the map* $jK\colon K \to S|K|$ *is a homotopy equivalence.*

It is also readily verified that

THEOREM (6.4). *Let* $K \in \mathcal{S}$ *and let* $f\colon S|K| \to S|K|$ *be a c.s.s. map such that commutativity holds in the diagram*

$$\begin{tikzcd} & S|K| \arrow[d, "f"] \\ K \arrow[ur, "jK"] \arrow[r, "jK"'] & S|K| \end{tikzcd}$$

Then f *is a homotopy equivalence.*

It follows from Corollary (6.2) and Theorems (6.3) and (6.4) that the pair $(S|\ |, j)$ is an H-pair. Application of Lemma (5.6) and Theorem (5.8) now yields

LEMMA (6.5). *Let* $K \in \mathcal{S}$. *Then the maps*

$$S|jK|\colon S|K| \to S|S|K\|, \qquad S|e^\infty K|\colon S|K| \to S|\mathrm{Ex}^\infty K|,$$

$$\mathrm{Ex}^\infty jK\colon \mathrm{Ex}^\infty K \to \mathrm{Ex}^\infty S|K|, \qquad \mathrm{Ex}^\infty e^\infty K\colon \mathrm{Ex}^\infty K \to \mathrm{Ex}^\infty \mathrm{Ex}^\infty K$$

are homotopy equivalences.

THEOREM (6.6). *Let* $N\colon \mathcal{S}_E \to \mathcal{Z}$ *be a homotopy notion on* $\mathcal{S}_E$. *Then the function* $h\colon N\,\mathrm{Ex}^\infty \to NS|\ |$ *given by*

$$hK = (NS|e^\infty K|)^{-1} \circ Nj\,\mathrm{Ex}^\infty K = (Ne^\infty S|K|)^{-1} \circ N\,\mathrm{Ex}^\infty jK$$

is a natural equivalence.

Theorem (6.6) asserts that the homotopy notions on $\mathcal{S}$ induced by the

functor Ex^∞ are equivalent with these induced by the functor $S|\ |$. In particular we have

COROLLARY (6.7). *Let $K \in \mathcal{S}$. Then* $\mathrm{Ex}^\infty K$ *and* $S|K|$ *have the same homotopy type.*

7. Extension and subdivision. The *subdivision* of a c.s.s. complex K is a c.s.s. complex $\mathrm{Sd}\,K$ defined as follows. Let $\bar{K}$ denote the c.s.s. complex of which the q-simplices are pairs (σ,ξ) such that $\sigma \in K$, $\xi \in \Delta'[\dim\sigma]$ and $\dim\xi = q$, while for a map $\gamma\colon [p] \to [q]$ the p-simplex $(\sigma,\xi)\gamma$ is given by $(\sigma,\xi)\gamma = (\sigma,\xi\gamma)$. Define a relation on $\bar{K}$ by calling two simplices $(\sigma,\xi), (\tau,\rho) \in \bar{K}$ equivalent if there exists a map $\alpha\colon [\dim\tau] \to [\dim\sigma]$ such that $\tau = \sigma\alpha$ and $\xi = \Delta'\alpha(\rho)$ and let $\sim$ denote the resulting equivalence relation. Then $\mathrm{Sd}\,K$ is the collapsed complex $\mathrm{Sd}\,K = \bar{K}/(\sim)$.

A c.s.s. map $f\colon K \to L$ clearly induces a c.s.s. map $\bar{f}\colon \bar{K} \to \bar{L}$ (given by $\bar{f}(\sigma,\xi) = (f\sigma,\xi)$) which is compatible with the relation $\sim$. The *subdivision* of f then is defined as the collapsed map $\mathrm{Sd}\,f\colon \mathrm{Sd}\,K \to \mathrm{Sd}\,L$. Clearly the function $\mathrm{Sd}\colon \mathcal{S} \to \mathcal{S}$ so defined is a covariant functor. By $\mathrm{Sd}^n\colon \mathcal{S} \to \mathcal{S}$ we shall mean the functor Sd applied n times.

The functors Ex and Sd are closely related. With a c.s.s. map f: $\mathrm{Sd}\,K \to L$ we may associate a c.s.s. map $\beta f\colon K \to \mathrm{Ex}\,L$ as follows. Let $\sigma \in K$ be an n-simplex and let $c\colon \bar{K} \to \mathrm{Sd}\,K$ be the collapsing map. Then $(\beta f)\sigma$ is the n-simplex of $\mathrm{Ex}\,L$, i.e., the c.s.s. map $(\beta f)c\colon \Delta'[n] \to L$, given by $((\beta f)\sigma)\xi = (f\circ c)(\sigma,\xi)$. The function β is natural, i.e., for every two maps $a\colon K' \to K$ and $b\colon L \to L'$

$$\beta(b \circ f \circ \mathrm{Sd}\,a) = \mathrm{Ex}\,b \circ \beta f \circ a.$$

An important property of the function β is

LEMMA (7.1). *Let $K, L \in \mathcal{S}$. Then the function β establishes a one-to-one correspondence between the c.s.s. maps* $\mathrm{Sd}\,K \to L$ *and the c.s.s. maps* $K \to \mathrm{Ex}\,L$.

Lemma (7.1) is an immediate consequence of the results of [7]. It can also be verified by a straightforward computation

For every c.s.s. complex K define an epimorphism $dK\colon K \to K$ as follows. Let $\bar{d}K\colon \bar{K} \to K$ be the map given by

$$\bar{d}K(\sigma,\xi) = (\phi_\sigma \circ \delta[\dim\sigma])\xi,$$

where $\phi_\sigma\colon \Delta[\dim\sigma] \to K$ is the (unique) map such that $\phi_\sigma\alpha = \sigma\alpha$ for all

$\alpha \in \Delta[\dim \sigma]$. Then $\bar{d}K$ maps equivalent simplices of $\bar{K}$ into the same simplex of K and dK: $\mathrm{Sd}\, K \to K$ is defined as the map obtained by collapsing $\bar{d}K$. By $d^n K$: $\mathrm{Sd}^n K \to K$ we shall mean the composite epimorphism

$$\mathrm{Sd}^n K \xrightarrow{d(\mathrm{Sd}^{n-1} K)} \mathrm{Sd}^{n-1} K \to \cdots \to \mathrm{Sd}\, K \xrightarrow{dK} K$$

It is readily verified that the function d is a natural transformation $d: Sd \to E$.

The natural transformations $e: E \to \mathrm{Ex}$ and d: $\mathrm{Sd} \to E$ are also closely related. In fact a simple computation yields

LEMMA (7.2). *Let $K \in \mathcal{S}$. Then $\beta(dK) = eK$.*

Remark (7.3). Lemma (7.1) states that, in the terminology of [6], the functor Sd is a left adjoint of the functor Ex.

Remark (7.4). The ordered sets $[n]$ and the maps α: $[m] \to [n]$ form a category which will be denoted by $\mathcal{V}$. The subdivided standard simplices $\Delta'[n]$ and the maps $\Delta'\alpha$: $\Delta'[m] \to \Delta'[n]$ now may be considered as the images of the objects $[n]$ and maps $\alpha: [m] \to [n]$ of the category $\mathcal{V}$ under a covariant functor Δ': $\mathcal{V} \to \mathcal{S}$. It then may be verified that the functors Sd and Ex may be obtained by the general method of [7], Section 3 by putting $\mathcal{Y} = \mathcal{S}$ and $\Sigma = \Delta'$.

Let $K \in \mathcal{S}$. A q-simplex of $\mathrm{Ex}^\infty K$ is a pair (σ, n) where $\sigma \in \mathrm{Ex}^n K$ is a q-simplex. As $\mathrm{Ex}^n K = \mathrm{Ex}^{n-1}(\mathrm{Ex}\, K)$ it follows that the pair $(\sigma, n-1)$ is a q-simplex of $\mathrm{Ex}^\infty(\mathrm{Ex}\, K)$. It is readily verified that this correspondence yields an isomorphism i: $\mathrm{Ex}^\infty K \to \mathrm{Ex}^\infty(\mathrm{Ex}\, K)$ such that commutativity holds in the diagram

$$\begin{array}{ccc} K & \xrightarrow{e^\infty K} & \mathrm{Ex}^\infty K \\ \downarrow{\scriptstyle eK} & & \downarrow{\scriptstyle i} \\ \mathrm{Ex}\, K & \xrightarrow{e^\infty(\mathrm{Ex}\, K)} & \mathrm{Ex}^\infty(\mathrm{Ex}\, K) \end{array} \tag{7.3}$$

In view of Lemma (6.5) the maps $S \mid e^\infty K \mid$ and $S \mid e^\infty(\mathrm{Ex} K) \mid$ are homotopy equivalences. Consequently the maps $\mid e^\infty K \mid$ and $\mid e^\infty(\mathrm{Ex} K) \mid$ are homotopy equivalences and it follows from the commutativity in diagram (7.3) that

LEMMA (7.4). *Let $K \in \mathcal{S}$. Then the continuous map $\mid eK \mid : \mid K \mid \to \mid \mathrm{Ex}\, K \mid$ is a homotopy equivalence.*

The following can be shown using standard methods.

Lemma (7.5). *Let $K \in \mathcal{S}$. Then the continuous map $| dK | : | \mathrm{Sd}\, K | \to | K |$ is a homotopy equivalence.*

8. C. s. s. approximation theorems. We shall now give an exact formulation of the c. s. s. approximation theorems mentioned in the introduction.

Theorem (8.1). *Let $K \in \mathcal{S}$ and let $M \in \mathcal{S}_E$. Then for every continuous map $f\colon | K | \to | M |$ there exists a c.s.s. map $h : K \to M$ such that $| h | \simeq f$.*

Let $L \in \mathcal{S}$ and let $M = \mathrm{Ex}^\infty L$. Then Theorem (8.1) implies

Corollary (8.2). *Let $K, L \in \mathcal{S}$. Then for every continuous map $f\colon | K | \to | L |$ there exists a c.s.s. map $h : K \to \mathrm{Ex}^\infty L$ such that the diagram*

$$
\begin{tikzcd}
{| K |} \arrow[r, "f"] \arrow[dr, "| h |"'] & {| L |} \arrow[d, "| e^\infty L |"] \\
& {| \mathrm{Ex}^\infty L |}
\end{tikzcd}
$$

is commutative up to homotopy, i. e., $| h | \simeq | e^\infty L | \circ f$.

Proof of Theorem (8.1). Let $\bar{j}M\colon S | M | \to M$ be a homotopy inverse of the map $jM\colon M \to S | M |$. Consider the diagram

$$
\begin{tikzcd}
{| K |} \arrow[r, "| jK |"] \arrow[d, "f"] & {| S | K ||} \arrow[d, "| Sf |"] & {| K |} \arrow[l, "| jK |"'] \arrow[d, "| h |"] \\
{| M |} \arrow[r, "| jM |"] & {| S | M ||} & {| M |} \arrow[l, "| jM |"']
\end{tikzcd}
$$

where $h : K \to M$ is the composite map

$$K \xrightarrow{jK} S | K | \xrightarrow{Sf} S | M | \xrightarrow{\bar{j}M} M.$$

Clearly commutativity holds in the rectangle at the left and the definition of h implies that the rectangle at the right is commutative up to homotopy. It follows from Lemma (6.6) that the maps $S | jK |$ and $S | jM |$ and therefore the maps $| jK |$ and $| jM |$ are homotopy equivalences. Hence $| h | \simeq f$.

A c. s. s. complex K is called *finite* if it has only a finite number of non-degenerate simplices.

THEOREM (8.3). *Let $K, L \in \mathcal{S}$ and let K be finite. Then for every continuous map $f: |K| \to |L|$ there exists an integer $n > 0$ and a c.s.s. map $h: K \to \mathrm{Ex}^n L$ such that the diagram*

$$\begin{tikzcd} {|K|} \arrow[r, "f"] \arrow[dr, "{|h|}"] & {|L|} \arrow[d, "{|e^n L|}"] \\ & {|\mathrm{Ex}^n L|} \end{tikzcd}$$

is commutative up to homotopy, i.e., $|h| \simeq |e^n L| \circ f$.

Proof. Application of Corollary (8.2) yields a c.s.s. map $h': K \to \mathrm{Ex}^\infty L$ such that $|h'| \simeq |e^\infty L| \circ f$. As K is finite only a finite number of non-degenerate simplices of $\mathrm{Ex}^\infty L$ are in the image of K under h'. Hence there exists an integer n such that the map $h': K \to \mathrm{Ex}^\infty L$ may be factorized

$$K \xrightarrow{h} \mathrm{Ex}^n L \xrightarrow{b} \mathrm{Ex}^\infty L$$

where b is the embedding map which assigns to a simplex $\sigma \in \mathrm{Ex}^n L$ the simplex $(\sigma, n) \in \mathrm{Ex}^\infty L$. By an argument similar to that used in the proof of Lemma (7.4) it follows that $|b|$ is a homotopy equivalence. The theorem now follows from the fact that the map $e^\infty L: L \to \mathrm{Ex}^\infty L$ may be factorized

$$L \xrightarrow{e^n L} \mathrm{Ex}^n L \xrightarrow{b} \mathrm{Ex}^\infty L.$$

In order to obtain the dual theorem, involving the functor Sd instead of Ex, we need the following lemma

LEMMA (8.4). *Let $K, L \in \mathcal{S}$. Then for every c.s.s. map $h: K \to \mathrm{Ex}\, L$ the diagram*

$$\begin{tikzcd} {|K|} \arrow[r, "{|h|}"] & {|\mathrm{Ex}\, L|} \\ {|\mathrm{Sd}\, K|} \arrow[u, "{|dK|}"] \arrow[r, "{|\beta^{-1}h|}"] & {|L|} \arrow[u, "{|eL|}"'] \end{tikzcd}$$

is commutative up to homotopy, i.e., $|eL| \circ |\beta^{-1}h| \simeq |h| \circ |dK|$.

The proof will be given in Section 16.

Applying Lemma (8.4) n times to Theorem (8.3) we get

THEOREM (8.5). *Let $K, L \in \mathcal{S}$ and let K be finite. Then for every continuous map $f\colon |K| \to |L|$ there exists an integer $n > 0$ and a c.s.s. map $g\colon Sd^n K \to L$ such that the diagram*

$$\begin{array}{ccc} |K| & \xrightarrow{\quad f \quad} & |L| \\ \uparrow {\scriptstyle |d^nK|} & \nearrow {\scriptstyle |g|} & \\ |\mathrm{Sd}^n K| & & \end{array}$$

is commutative up to homotopy, i.e., $|g| \simeq f \circ |d^nK|$.

Chapter II. Proofs.

9. Proof of Lemma (3.1). Let $f_0, f_1\colon K \to L \in \mathcal{S}$ be maps such that $f_0 \simeq f_1$. Using the terminology of [4] this means that there exists a c.s.s. map $f_I\colon I \times K \to L$ such that $f_I \circ \epsilon K = f_\epsilon$ $(\epsilon = 0, 1)$. It is readily verified that the functor Ex commutes with the cartesian product, i.e., that for every two c.s.s. complexes A and B

$$\mathrm{Ex}(A \times B) = (\mathrm{Ex}\ A) \times (\mathrm{Ex}\, B).$$

Straightforward computation shows that commutativity holds in the diagram

$$\begin{array}{ccc} \mathrm{Ex}\, K & \xrightarrow{\ \epsilon(\mathrm{Ex}\, K)\ } & I \times (\mathrm{Ex}\, K) \\ \downarrow {\scriptstyle \mathrm{Ex}(\epsilon K)} & & \downarrow {\scriptstyle eI \times i_{\mathrm{Ex}\, K}} \\ \mathrm{Ex}(I \times K) & \xrightarrow{\quad i \quad} & (\mathrm{Ex}\, I) \times (\mathrm{Ex}\, K) \end{array}$$

where i is the identity. Hence

$$(\mathrm{Ex}\, f_I) \circ (eI \times i_{\mathrm{Ex}\, K}) \circ \epsilon(\mathrm{Ex}\, K) = (\mathrm{Ex}\, f_I) \circ (\mathrm{Ex}(\epsilon K)) = \mathrm{Ex}(f_I \circ \epsilon K) = \mathrm{Ex}\, f_\epsilon,$$

i.e., $(\mathrm{Ex}\, f_I) \circ (eI \times i_{\mathrm{Ex}\, K})\colon \mathrm{Ex}\, f_0 \simeq \mathrm{Ex}\, f_1$.

10. Proof of Lemma (3.2). We shall first investigate the structure of $\mathrm{Ex}\, K$.

A map $\alpha\colon [m] \to [n]$ was defined as a monotone function. By a *function* $\zeta\colon [m] \to [n]$ we shall mean merely a function which thus need not be monotone. A *permutation* $\pi\colon [m] \to [m]$ is a function which is one-to-one onto.

Let $\pi: [m] \to [m]$ be a permutation. Then π induces an automorphism $\pi': \Delta'[m] \to \Delta'[m]$ as follows. For each map $\sigma: [q] \to [m]$ let $\sigma^\pi: [q] \to [m]$ be a map and let $\phi: [q] \to [q]$ be a permutation such that commutativity holds in the diagram

$$\begin{tikzcd} {[q]} \arrow[r, "\sigma"] \arrow[d, "\phi"] & {[m]} \arrow[d, "\pi"] \\ {[q]} \arrow[r, "\sigma^\pi"] & {[m]} \end{tikzcd}$$

Clearly such a map σ^π and permutation ϕ exist. It is easily seen that

(a) σ^π is unique;

(b) if σ is a monomorphism, then so is σ^π;

(c) if σ lies on τ, then σ^π lies on τ^π.

We now define the automorphism $\pi': \Delta'[m] \to \Delta'[m]$ by

$$\pi'(\sigma_0, \cdot\cdot\cdot, \sigma_q) = (\sigma_0{}^\pi, \cdot\cdot\cdot, \sigma_q{}^\pi).$$

Let $\zeta: [m] \to [n]$ be a function. Then ζ induces a c. s. s. map ζ': $\Delta'[m] \to \Delta'[n]$ as follows. There clearly exists a permutation $\pi: [m] \to [m]$ and a unique map $\alpha: [m] \to [n]$ such that commutativity holds in the diagram

$$\begin{tikzcd} {[m]} \arrow[d, "\pi"] \arrow[dr, "\zeta"] & \\ {[m]} \arrow[r, "\alpha"] & {[n]} \end{tikzcd}$$

The c. s. s. map $\zeta': \Delta'[m] \to \Delta'[n]$ is now defined as the composite map

$$\Delta'[m] \xrightarrow{\pi'} \Delta'[m] \xrightarrow{\Delta'\alpha} \Delta'[n].$$

It is readily verified that

(a) the c. s. s. map ζ' is independent of the choice of the permutation π;

(b) if ζ is a permutation, then this definition of ζ' coincides with the above one;

(c) if ζ is a map, then $\zeta' = \Delta'\zeta$;

(d) if $\vartheta: [l] \to [m]$ is a function, then $(\zeta\vartheta)'$ is the composite map;

$$\Delta'[l] \xrightarrow{\vartheta'} \Delta'[m] \xrightarrow{\zeta'} \Delta'[n].$$

Ex K is a c. s. s. complex. This means that for every n-simplex $\sigma \in \operatorname{Ex} K$ and every *map* $\alpha : [m] \to [n]$ there is given an m-simplex $\sigma\alpha \in \operatorname{Ex} K$ such that

(i) $\sigma\epsilon_n = \sigma$ where $\epsilon_n : [n] \to [n]$ is the identity;

(ii) if $\beta : [l] \to [m]$ is a map, then $(\sigma\alpha)\beta = \sigma(\alpha\beta)$.

Now let $\sigma \in \operatorname{Ex} K$ be an n-simplex and let $\zeta : [m] \to [n]$ be a *function.* Then the composite map

$$\Delta'[m] \xrightarrow{\zeta'} \Delta'[n] \xrightarrow{\sigma} K$$

is an m-simplex of Ex K which will be denoted by $\sigma\zeta$. If $\vartheta : [l] \to [m]$ is also a function, then clearly $(\sigma\zeta)\vartheta = \sigma(\zeta\vartheta)$. Thus Ex K *has more structure than a c. s. s. complex.* It is this additional structure which will be used in the proof of Lemma (3.2).

Proof of Lemma (3.2). Let $\Lambda \subset \Delta[n]$ be the subcomplex generated by the non-degenerate $(n-1)$-simplices $\epsilon^0, \cdots, \epsilon^{k-1}, \epsilon^{k+1}, \cdots, \epsilon^n$ and let $\lambda : \Lambda \to \operatorname{Ex} K$ be the c. s. s. map such that $\lambda\epsilon^i = \tau_i$. Then we must define a c. s. s. map $\rho : \Delta'[n] \to \operatorname{Ex} K$ such that for each $i \neq k$ commutativity holds in the diagram

(10.1)
$$\begin{tikzcd}
\Delta'[n-1] \arrow[r, "\Delta'\epsilon^i"] \arrow[d, "{\delta[n-1]}"] & \Delta'[n] \arrow[dr, "\rho"] & \\
 & & \operatorname{Ex} K \\
\Delta[n-1] \arrow[r, "\Delta\epsilon^i"] & \Lambda \arrow[ur, "\lambda"'] &
\end{tikzcd}$$

For each simplex $(\sigma_0, \cdots, \sigma_q) \in \Delta'[n]$ define a function $\zeta(\sigma_0, \cdots, \sigma_q)$: $[q] \to [n]$ by

$$\zeta(\sigma_0, \cdots, \sigma_q)(i) = \sigma_i(\dim \sigma_i), \qquad \sigma_i \neq \epsilon^k \text{ or } \epsilon_n$$

$$\zeta(\sigma_0, \cdots, \sigma_q)(i) = k, \qquad \sigma_i = \epsilon^k \text{ or } \epsilon_n.$$

Then there exists a permutation $\phi : [q] \to [q]$ and a unique map $\sigma : [q] \to [n]$ such that commutativity holds in the diagram

$$\begin{tikzcd}
{[q]} \arrow[d, "\phi"] \arrow[dr, "{\zeta(\sigma_0, \cdots, \sigma_q)}"] & \\
{[q]} \arrow[r, "\sigma"] & {[n]}
\end{tikzcd}$$

It is easily seen that $\sigma \in \Lambda$. We now define $\rho(\sigma_0, \cdots, \sigma_q) = (\lambda\sigma)\phi$. It may be verified by direct computation that this definition is independent of the choice of the permutation ϕ.

We now show that the function $\rho\colon \Delta'[n] \to \operatorname{Ex} K$ so defined is a c. s. s. map. Let $\beta\colon [p] \to [q]$ be a map. Then there exists a permutation $\psi\colon [p] \to [p]$ and a unique map $\gamma\colon [p] \to [q]$ such that commutativity holds in the diagram

$$\begin{tikzcd}
{[p]} \arrow[r, "\beta"] \arrow[d, "\psi"] & {[q]} \arrow[d, "\phi"] \arrow[dr, "\zeta(\sigma_0, \cdots, \sigma_q)"] & \\
{[p]} \arrow[r, "\gamma"] & {[q]} \arrow[r, "\sigma"] & {[n]}
\end{tikzcd}$$

The function $\zeta((\sigma_0, \cdots, \sigma_q)\beta)$ is the composite function

$$[p] \xrightarrow{\beta} [q] \xrightarrow{\zeta(\sigma_0, \cdots, \sigma_q)} [n]$$

and consequently $\rho((\sigma_0, \cdots, \sigma_q)\beta) = (\lambda(\sigma\gamma))\psi$. As commutativity also holds in the diagram

$$\begin{tikzcd}
\Delta'[p] \arrow[r, "\Delta'\beta"] \arrow[d, "\psi'"] & \Delta'[q] \arrow[d, "\phi'"] & \\
\Delta'[p] \arrow[r, "\Delta'\gamma"] & \Delta'[q] \arrow[r, "\lambda\sigma"] & K
\end{tikzcd}$$

it follows that

$$\lambda(\sigma\gamma))\psi = \lambda\sigma \circ \Delta'\gamma \circ \psi' = \lambda\sigma \circ \phi' \circ \Delta'\beta = ((\lambda\sigma)\pi)\beta$$

i. e., the function $\rho\colon \Delta'[n] \to \operatorname{Ex} K$ is a c. s. s. map.

It thus remains to show that commutativity holds in diagram (10.1). Let $(\tau_0, \cdots, \tau_q) \in \Delta'[n-1]$. Then

$$\Delta'\epsilon^i(\tau_0, \cdots, \tau_q) = (\epsilon^i\tau_0, \cdots, \epsilon^i\tau_q).$$

If $i \neq k$, then clearly $\epsilon^i\tau_j \neq \epsilon^k$ and $\epsilon^i\tau_j \neq \epsilon_n$ for all j and it follows from the definitions of the maps ρ and $\delta[n]$ that

$$(\rho \circ \Delta'\epsilon^i)(\tau_0, \cdots, \tau_q) = (\lambda \circ \delta[n] \circ \Delta'\epsilon^i)(\tau_0, \cdots, \tau_q).$$

Application of Lemma (2.2) now yields

$$(\rho \circ \Delta'\epsilon^i)(\tau_0, \cdots, \tau_q) = (\lambda \circ \Delta\epsilon^i \circ \delta[n-1])(\tau_0, \cdots, \tau_q).$$

This completes the proof.

11. Proof of Lemma (3.4). Let k be an integer with $0 \leqq k \leqq n$, let $\tau_0, \cdots, \tau_{k-1}, \tau_{k+1}, \cdots, \tau_n \in \operatorname{Ex} K$ be n $(n-1)$-simplices such that $\tau_i\epsilon^{j-1} = \tau_j\epsilon^i$ for $i < j$ and $i \neq k \neq j$ and let $\rho \in \operatorname{Ex} L$ be an n-simplex such that $(\operatorname{Ex} f)\tau_i = \rho\epsilon^i$ for $i = 0, \cdots, \hat{k}, \cdots, n$. Then in order to prove the first part of Lemma (3.4) we must show that there exists a c.s.s. map $\tau\colon \Delta'[n] \to K$ such that for each integer $i \neq k$ commutativity holds in the diagram

(11.1)
$$\begin{tikzcd}
\Delta'[n-1] \arrow[r, "\tau_i"] \arrow[d, "\Delta'\epsilon^i"'] & K \arrow[d, "f"] \\
\Delta'[n] \arrow[r, "\rho"'] \arrow[ur, "\tau"] & L
\end{tikzcd}$$

For each simplex $(\sigma_0, \cdots, \sigma_q) \in \Delta'[n]$ for which there exists an integer $i \neq k$ and a simplex $(\sigma_0{}^i, \cdots, \sigma_q{}^i) \in \Delta'[n-1]$ such that $\Delta'\epsilon^i(\sigma_0{}^i, \cdots, \sigma_q{}^i) = (\sigma_0, \cdots, \sigma_q)$ define

$$\tau(\sigma_0, \cdots, \sigma_q) = \tau_i(\sigma_0{}^i, \cdots, \sigma_q{}^i).$$

This definition is independent of the choice of i. If j is another such integer and $i < j$ then there exists a simplex $(\sigma_0{}^{ij}, \cdots, \sigma_q{}^{ij}) \in \Delta'[n-2]$ such that $\Delta'\epsilon^{j-1}(\sigma_0{}^{ij}, \cdots, \sigma_q{}^{ij}) = (\sigma_0{}^i, \cdots, \sigma_q{}^i)$ and $\Delta'\epsilon^i(\sigma_0{}^{ij}, \cdots, \sigma_q{}^{ij}) = (\sigma_0{}^j, \cdots, \sigma_q{}^j)$.

Hence

$$\tau_i(\sigma_0{}^i, \cdots, \sigma_q{}^i) = \tau_i(\Delta'\epsilon^{j-1}(\sigma_0{}^{ij}, \cdots, \sigma_q{}^{ij})) = \tau_i\epsilon^{j-1}(\sigma_0{}^{ij}, \cdots, \sigma_q{}^{ij})$$
$$= \tau_j\epsilon^i(\sigma_0{}^{ij}, \cdots, \sigma_q{}^{ij}) = \tau_j(\Delta'\epsilon^i(\sigma_0{}^{ij}, \cdots, \sigma_q{}^{ij})) = \tau_j(\sigma_0{}^j, \cdots, \sigma_q{}^j).$$

It is readily verified that the function τ so defined on all simplices of $\Delta'[n]$ which are in the image of $\Delta'[n-1]$ under a map $\Delta'\epsilon^i$ with $i \neq k$, (i.e., those simplices $(\sigma_0, \cdots, \sigma_q) \in \Delta'[n]$ for which $\sigma_q \neq \epsilon_n$ or ϵ^k), commutes with all operators $\beta\colon [p] \to [q]$ and is such that commutativity holds in the upper left triangle of diagram (11.1).

It thus remains to show that τ can be extended over all of $\Delta'[n]$ (i.e., over the simplices $(\sigma_0, \cdots, \sigma_q) \in \Delta'[n]$ for which $\sigma_q = \epsilon_n$ or ϵ^k) to a c.s.s.

map in such a manner that commutativity also holds in the lower right triangle of diagram (11.1). For each non-degenerate simplex $(\sigma_0, \cdots, \sigma_q)$ with $\sigma_q = \epsilon_n$ let $T(\sigma_0, \cdots, \sigma_q)$ denote the triple (l, m, q) where l is the smallest integer such that $\sigma_l(i) = k$ for some i and $m = \dim \sigma_l$. Order these triples lexicographically. It is readily verified that

(i) if $T(\sigma_0, \cdots, \sigma_q) = (l, m, q)$ and $\dim_{l-1} < m - 1$ or $l = 0, m > 0$, then there exists a simplex $(\sigma_0', \cdots, \sigma_{q+1}') \in \Delta'[n]$ such that $(\sigma_0', \cdots, \sigma_{q+1})\epsilon^l = (\sigma_0, \cdots, \sigma_q)$ and $T(\sigma_0', \cdots, \sigma_{q+1}') = (l, m-1, q+1) < (l, m, q)$.

(ii) if $T(\sigma_0, \cdots, \sigma_q) = (l, m, q)$ and $\dim \sigma_{l-1} = m - 1, l < q$ or $l = m = 0$, then (a) $T((\sigma_0, \cdots, \sigma_q)\epsilon^i) < (l, m, q)$ for $i \neq l, q$, (b) $\sigma_{q-1} \neq \epsilon^k$ and hence $\tau((\sigma_0, \cdots, \sigma_q)\epsilon^q)$ has already been defined, (c) $T((\sigma_0, \cdots, \sigma_q)\epsilon^l) > (l, m, q)$ and (d) if $T(\sigma_0', \cdots, \sigma_q') \leqq (l, m, q)$, then $(\sigma_0, \cdots, \sigma_q)\epsilon^l$ is not a face of $(\sigma_0', \cdots, \sigma_q')$.

(iii) if $T(\sigma_0, \cdots, \sigma_q) = (q, n, q)$ and $\dim \sigma_{l-1} = n - 1$, then (a) $T((\sigma_0, \cdots, \sigma_q)\epsilon^i) < (q, n, q)$ for $i \neq q$, (b) $\sigma_{q-1} = \epsilon^k$ and (c) if $T(\sigma_0', \cdots, \sigma_q') \leqq (q, n, q)$, then $(\sigma_0, \cdots, \sigma_{q-1})$ is not a face of $(\sigma_0', \cdots, \sigma_q')$.

We now extend τ as follows. Let (l, m, q) be a triple and suppose that τ has already been extended over all non-degenerate simplices $(\sigma_0, \cdots, \sigma_{q-1}, \epsilon_n)$ and their faces for which $T(\sigma_0, \cdots, \sigma_{q-1}, \epsilon_n) < (l, m, q)$ and over some non-degenerate simplices $(\sigma_0, \cdots, \sigma_{q-1}, \epsilon_n)$ and their faces for which $T(\sigma_0, \cdots, \sigma_{q-1}, \epsilon_n) = (l, m, q)$ in such a manner that τ commutes with all face operators and that commutativity holds in the lower right triangle of diagram (11.1). Let $(\sigma_0, \cdots, \sigma_{q-1}, \epsilon_n)$ be a non-degenerate simplex such that $T(\sigma_0, \cdots, \sigma_{q-1}, \epsilon_n) = (l, m, q)$ and on which τ has not yet been defined. It then follows from (i) that $\dim \sigma_{l-1} = m - 1$ or $l = m = 0$ and from (ii) or (iii) that τ already has been defined on all faces of $(\sigma_0, \cdots, \sigma_{q-1}, \epsilon_n)$ except $(\sigma_0, \cdots, \sigma_{q-1}, \epsilon_n)\epsilon^l$. Because f is a fibre map there exists a q-simplex $\psi \in K$ such that

$$\rho(\sigma_0, \cdots, \sigma_{q-1}, \epsilon_n) = f\psi, \qquad \tau((\sigma_0, \cdots, \sigma_{q-1}, \epsilon_n)\epsilon^i) = \psi\epsilon^i \qquad (i \neq k).$$

Now define

$$\tau(\sigma_0, \cdots, \sigma_{q-1}, \epsilon_n) = \psi, \qquad \tau((\sigma_0, \cdots, \sigma_{q-1}, \epsilon_n)\epsilon^l) = \psi\epsilon^l.$$

It is readily verified that the function τ so extended commutes with all face operators and is such that commutativity holds in the lower right triangle of diagram (11.1). Thus using induction on the triples (l, m, q) τ may be extended over all non-degenerate simplices $(\sigma_0, \cdots, \sigma_{q-1}, \epsilon_n) \in \Delta'[n]$ and their faces. As every non-degenerate simplex $(\sigma_0, \cdots, \sigma_{q-2}, \epsilon^k) \in \Delta'[n]$ is a face

of a non-degenerate simplex $(\sigma_0, \cdots, \sigma_{q-2}, \epsilon^k, \epsilon_n)$ it follows that τ may be extended over all non-degenerate simplices of $\Delta'[n]$ in such a manner that τ commutes with all face operators and that commutativity holds in diagram (11.1). Extensions of τ over the degenerate simplices of $\Delta'[n]$ (which is always possible in a unique way) now yields the desired c. s. s. map τ: $\Delta'[n] \to K$.

The second part of Lemma (3.4) is obvious.

12. Proof of Lemma (3.6). We shall use the theory of acyclic models of Eilenberg-MacLane (see [1]). The models will be the complexes $\Delta[n]$ and $\Delta'[n]$. Let C_a: $\mathcal{S} \to \partial\mathcal{G}$ be the augmented chain functor. As the map $eK: K \to \operatorname{Ex} K$ induces a one-to-one correspondence between the 0-simplices of K and those of $\operatorname{Ex} K$ it is sufficient to prove that

(a) the functor C_a: $\mathcal{S} \to \partial\mathcal{G}$ is representable in dimension > 0,

(b) the composite functor $\mathcal{S} \xrightarrow{\operatorname{Ex}} \mathcal{S} \xrightarrow{C_a} \partial\mathcal{G}$ is representable in dimension > 0, and

(c) for every integer $n \geqq 0$,

$$H_*(\Delta[n]) = H_*(\operatorname{Ex}\Delta[n]) = 0, \qquad H_*(\Delta'[n]) = H_*(\operatorname{Ex}\Delta'[n]) = 0.$$

Let $K \in \mathcal{S}$, for every n-simplex $\sigma \in K$ let ϕ_σ: $\Delta[n] \to K$ be the unique c. s. s. map such that $\phi_\sigma \alpha = \sigma\alpha$ for all $\alpha \in \Delta[n]$ and let ϵ_n' be the generator of $C_a\Delta[n]$ corresponding to the identity map ϵ_n: $[n] \to [n]$, i. e., the only non-degenerate n-simplex of $\Delta[n]$. Then it is easily seen that the function $\sigma \to (\sigma, \epsilon_n')$ yields a representation of the functor C_a.

Let $K \in \mathcal{S}$, let τ: $\Delta'[n] \to K$ be an n-simplex of $\operatorname{Ex} K$ and let ι_n' be the generator of $C_a\operatorname{Ex}\Delta'[n]$ corresponding with the identity map $\iota_n: \Delta'[n] \to \Delta'[n]$. Then it is easily seen that the function $\tau \to (\tau, \iota_n')$ yields a representation of the functor $C_a\operatorname{Ex}$.

For every integer $n \geqq 0$ the (unique) map $\Delta[n] \to \Delta[0]$ is a homotopy equivalence in $\mathcal{S}$. Combining this with Lemma (3.1) and the fact that $\Delta[0] \approx \operatorname{Ex}\Delta[0]$ and $H_*(\Delta[0]) = 0$ we get $H_*(\Delta[n]) = H_*(\operatorname{Ex}\Delta[n]) = 0$. If for each integer $n \geqq 0$ the map $\delta[n]: \Delta'[n] \to \Delta[n]$ is a homotopy equivalence, then $H_*(\Delta'[n]) = H_*(\Delta[n]) = 0$, and Lemma (3.1) implies $H_*(\operatorname{Ex}\Delta'[n]) = H_*(\operatorname{Ex}\Delta[n]) = 0$. It thus remains to show that $\delta[n]$ is a homotopy equivalence.

For each integer i with $0 \leqq i \leqq n$ let $\beta_i: [i] \to [n]$ be the map given by $\beta_i(j) = j$, $0 \leqq j \leqq i$. Define a function $\delta'[n]: \Delta[n] \to \Delta'[n]$ by $\delta[n]\sigma = (\beta_{\sigma(0)}, \cdots, \beta_{\sigma(q)})$, $\dim \sigma = q$. As for every map $\alpha: [p] \to [q]$,

$$(\delta'[n]\sigma)\alpha = (\beta_{\sigma(0)}, \cdots, \beta_{\sigma(q)})\alpha = (\beta_{\sigma\alpha(0)}, \cdots, \beta_{\sigma\alpha(p)}) = \delta'[n](\sigma\alpha),$$

it follows that $\delta'[n]$ is a c.s.s. map. The composite map

$$\Delta[n] \xrightarrow{\delta'[n]} \Delta'[n] \xrightarrow{\delta[n]} \Delta[n]$$

is the identity because for $\sigma \in \Delta[n]$ and $0 \leqq i \leqq \dim \sigma$

$$(\delta[n]\delta'[n]\sigma)(i) = \beta_{\sigma(i)}(\sigma(i)) = \sigma(i).$$

It thus remains to prove that the composite map

$$\Delta'[n] \xrightarrow{\delta[n]} \Delta[n] \xrightarrow{\delta'[n]} \Delta'[n]$$

is homotopic with the identity $\iota_n \colon \Delta'[n] \to \Delta'[n]$.

For each simplex $\sigma \in \Delta[n]$, let $\bar{\sigma} = \beta_{\sigma(\dim \sigma)}$. Define a function h: $\Delta[1] \times \Delta'[n] \to \Delta'[n]$ by

$$h(\epsilon^0\eta^0 \cdots \eta^{q-1}, (\sigma_0, \cdots, \sigma_q)) = (\bar{\sigma}_0, \cdots, \bar{\sigma}_q),$$

$$h(\epsilon^1\eta^0 \cdots \eta^{q-1}, (\sigma_0, \cdots, \sigma_q)) = (\sigma_0, \cdots, \sigma_q),$$

$$h(\epsilon_1\eta^0 \cdots \eta^{i-1}\eta^{i+1} \cdots \eta^{q-1}, (\sigma_0, \cdots, \sigma_q)) = (\sigma_0, \cdots, \sigma_i, \bar{\sigma}_{i+1}, \cdots, \bar{\sigma}_q).$$

A straightforward computation shows that the function h so defined is a c.s.s. map. It is now easily verified that h is the required homotopy.

13. Proof of Lemma (3.7). Use will be made of the following c.s.s. analogues of two theorems of J. H. C. Whitehead ([10]).

THEOREM (13.1). *Let $K, L \in \mathcal{S}_E$ be connected and let $\phi \in K$ be a 0-simplex. Then a c.s.s. map $f \colon K \to L$ is a homotopy equivalence if and only if f induces isomorphisms of all homotopy groups, i.e., $f_* \colon \pi_n(K;\phi) \approx \pi_n(L; f\phi)$, $n \geqq 1$.*

THEOREM (13.2). *Let $K, L \in \mathcal{S}_E$ be simply connected. Then a c.s.s. map $f \colon K \to L$ is a homotopy equivalence if and only if f induces isomorphisms of all homology groups, i.e., $f_* \colon H_*(K) \approx H_*(L)$.*

We also need the following lemma

LEMMA (13.3). *Let $K \in \mathcal{S}_E$ and let $\phi \in K$ be a 0-simplex. Then $(eK)_* \colon \pi_1(K;\phi) \approx \pi_1(\mathrm{Ex}\, K; (eK)\phi)$.*

Proof of Lemma (3.7). In this proof we shall freely use the results of [9] Clearly K may be supposed to be minimal. Let $\pi = \pi_1(K)$. Then

there exists a fibre map $p: K \to K(\pi, 1)$ with simply connected fibre F. Let $q: F \to K$ be the inclusion map, then it follows from the naturality of e that commutativity holds in the diagram

$$\begin{array}{ccccc} F & \xrightarrow{q} & K & \xrightarrow{p} & K(\pi,1) \\ \downarrow{\scriptstyle eF} & & \downarrow{\scriptstyle eK} & & \downarrow{\scriptstyle e(K(\pi,1))} \\ \operatorname{Ex} F & \xrightarrow{\operatorname{Ex} q} & \operatorname{Ex} K & \xrightarrow{\operatorname{Ex} p} & \operatorname{Ex} K(\pi,1) \end{array}$$

By Lemma (3.4) $\operatorname{Ex} p$ is a fibre map with $\operatorname{Ex} F$ as a fibre. Hence in order to prove that eK is a homotopy equivalence it is, in view of the exactness of the homotopy sequence of a fibre map, the "five lemma" and Theorem (13.1), sufficient to prove that eF and $e(K(\pi, 1))$ are homotopy equivalences.

As F is simply connected, so is $\operatorname{Ex} F$ (Lemma (13.3)). Hence it follows from Lemma (3.6) and Theorem (13.2) that eF is a homotopy equivalence.

There exists a fibre map $t: W(K(\pi, 0)) \to K(\pi, 1)$ with $K(\pi, 0)$ as fibre and, as above, in order to prove that $e(K(\pi, 1))$ is a homotopy equivalence it suffices to prove that $e(W(K(\pi, 0)))$ and $e(K(\pi, 0))$ are so. As $W(K(\pi, 0))$ is contractible and a fortiori simply connected the argument applied to F yields that $e(W(K(\pi, 0)))$ is a homotopy equivalence. It is also readily verified that $e(K(\pi, 0))$ is an isomorphism. Hence $e(K(\pi, 1))$ is a homotopy equivalence.

This completes the proof of Lemma (3.7).

Proof of Lemma (13.3). For a definition of the fundamental group see [9].

Let $\sigma \in \Delta[n]$ be a non-degenerate q-simplex, i.e., the map $\sigma: [q] \to [n]$ is a monomorphism. Then σ is completely determined by the set $(\sigma(0), \cdots, \sigma(q))$, the image of $[q]$ under σ. We shall often write $(\sigma(0), \cdots, \sigma(q))$ instead of σ.

We first prove that $(eK)_*: \pi_1(K; \phi) \to \pi_1(\operatorname{Ex} K; (eK)\phi)$ is a monomorphism. Let $a \in \pi_1(K; \phi)$ be such that $(eK)_* a = 1$ and let $\tau \in a$. Then there exists a 2-simplex $\rho \in \operatorname{Ex} K$ such that $\rho\epsilon^2 = (eK)\tau$ and $\rho\epsilon^0 = \rho\epsilon^1 = (eK)\phi\eta^0$. Iterated application of the extension condition yields 4 3-simplices $\tau_1, \tau_3, \tau_2, \tau_4 \in K$ such that

$$\tau_1\epsilon^1 = \rho((1), (0,1), (0,1,2)); \quad \tau_1\epsilon^2 = \rho((1), (1,2), (0,1,2)); \quad \tau_1\epsilon^3 = \phi\eta^0\eta^0$$

$$\tau_2\epsilon^0 = \tau_1\epsilon^0; \quad \tau_2\epsilon^2 = \rho((2), (1,2), (0,1,2)); \quad \tau_2\epsilon^3 = \phi\eta^0\eta^0$$

$$\tau_3\epsilon^1 = \tau_2\epsilon^1; \quad \tau_3\epsilon^2 = \rho((2),(0,2),(0,1,2)); \quad \tau_3\epsilon^3 = \phi\eta^0\eta^0$$

$$\tau_4\epsilon^0 = \tau_3\epsilon^0; \quad \tau_4\epsilon^1 = \rho((0),(0,1),(0,1,2)); \quad \tau_4\epsilon^2 = \rho((0),(0,2),(0,1,2)).$$

Then

$$\tau_4\epsilon^3\epsilon^0 = \tau_4\epsilon^0\epsilon^2 = \tau_3\epsilon^0\epsilon^2 = \tau_3\epsilon^3\epsilon^0 = \phi\eta^0$$

$$\tau_4\epsilon^3\epsilon^1 = \tau_4\epsilon^0\epsilon^2 = \rho((0),(0,1)) = \sigma$$

$$\tau_4\epsilon^3\epsilon^2 = \tau_4\epsilon^2\epsilon^2 = \rho((0),(0,2)) = \phi\eta^0.$$

Consequently $a = 1$.

We now show that $(ek)_*: \pi_1(K;\phi) \rightarrow \pi_1(\mathrm{Ex}\,K;(eK)\phi)$ is an epimorphism. Let $\psi \in b \in \pi_1(\mathrm{Ex}\,K;(eK)\phi)$. Define a c. s. s. map $\rho: \Delta'[2] \rightarrow K$ by $\rho((0),(0,1)) = \psi((0),(0,1))$, $\rho((1),(0,1)) = \psi((1),(0,1))$,

$$\rho((1),(1,2),(0,1,2)) = \rho((2),(0,2),(0,1,2)) = \rho((2),(1,2),(0,1,2)) = \phi\eta^0\eta^0,$$

and extend ρ over $((0),(0,1),(0,1,2))$, $((0),(0,2),(0,1,2))$ and $((1),(0,1),(0,1,2))$ by iterated application of the extension condition. Then

$$\rho\epsilon^0 = (eK)\phi\eta^0, \qquad \rho\epsilon^1 = (eK)\rho((0),(0,2)), \qquad \rho\epsilon^2 = \tau$$

Consequently there exists an element $a \in \pi_1(K,\phi)$ such that $\rho((0),(0,2)) \in a$ and $(eK)_*a = b$.

14. Proof of Theorem (4.6). Clearly K may suppose to be connected. Let $\phi \in \mathrm{Ex}^\infty K$ be a 0-simplex, then in view of Theorem (13.1) it suffices to prove that $f_*: \pi_n(\mathrm{Ex}^\infty K;\phi) \approx \pi_n(\mathrm{Ex}^\infty K; f\phi)$ for all $n \geqq 1$. We shall only give a proof for $n = 1$. The proof for $n > 1$ is similar although more complicated.

Let $a \in \pi_1(\mathrm{Ex}^\infty K;\phi)$ and let τ be a representant of a. Suppose there exists a 2-simplex $\rho \in \mathrm{Ex}^\infty K$ such that

$$(14.1) \qquad \rho\epsilon^0 = \tau\epsilon^0\eta^0, \qquad \rho\epsilon^1 = \tau, \qquad \rho\epsilon^2 = f\tau.$$

Then clearly $f_*a = a$. Hence it suffices to show that for every 1-simplex $\tau \in \mathrm{Ex}^\infty K$ there exists a 2-simplex $\rho \in \mathrm{Ex}^\infty K$ satisfying condition (14.1).

Let $\tau \in \mathrm{Ex}^\infty K$ be a 1-simplex and let n be the smallest integer $n \geqq 0$ such that $\tau = (\psi, n)$ (by $\tau = (\psi, 0)$ we mean $\tau = (e^\infty K)\psi$). If $n = 0$, then by hypothesis $\rho = \tau\eta^1$ is the desired 2-simplex. Now suppose it has already been proved that if $n < m$, then there exists a 2-simplex ρ satisfying (14.1a). Then we must show that this is also the case if $n = m$.

Define, using the notation of Section 13, a 2-simplex $\vartheta \in \mathrm{Ex}^n K$ as follows.

$$\vartheta((0),(0,1),(0,1,2))=\vartheta((0),(0,2),(0,1,2))=\psi((0),(0,1))\eta^1$$
$$\vartheta((1),(0,1),(0,1,2))=\vartheta((1),(1,2),(0,1,2))=\psi((1),(0,1))\eta^1$$
$$\vartheta((2),(0,2),(0,1,2))=\vartheta((2),(1,2),(0,1,2))=\psi((0,1))\eta^0\eta^1.$$

Then it is readily verified that

$$\vartheta\epsilon^0=(e(\mathrm{Ex}^{n-1}K))\psi((1),(0,1)),\qquad \vartheta\epsilon^1=(e(\mathrm{Ex}^{n-1}K))\psi((0),(0,1)),\qquad \vartheta\epsilon^2=\psi.$$

By the induction hypothesis there exist 2-simplices $\rho_0,\rho_1\in\mathrm{Ex}^\infty K$ such that

$$\rho_0\epsilon^0=(\psi((0,1))\eta^0,n-1),\qquad \rho_0\epsilon^1=(\vartheta\epsilon^0,n),\qquad \rho_0\epsilon^2=f(\vartheta\epsilon^0,n)$$
$$\rho_1\epsilon^0=(\psi((0,1))\eta^0,n-1),\qquad \rho_1\epsilon^1=(\vartheta\epsilon^1,n),\qquad \rho_1\epsilon^2=f(\vartheta\epsilon^1,n).$$

Application of the extension condition then yields 3-simplices $\kappa,\lambda\in\mathrm{Ex}^\infty K$ such that

$$\kappa\epsilon^0=\rho_0,\quad \kappa\epsilon^1=\rho_1,\quad \kappa\epsilon^3=f(\vartheta,n),$$
$$\lambda\epsilon^0=(\vartheta\epsilon^0\eta^0,n),\quad \lambda\epsilon^1=(\vartheta,n),\quad \lambda\epsilon^2=\kappa\epsilon^2.$$

It then follows by direct computation that $\lambda\epsilon^3$ is the desired 2-simplex, i.e.,

$$\lambda\epsilon^3\epsilon^0=\tau\epsilon^0\eta^0,\quad \lambda\epsilon^3\epsilon^1=\tau,\quad \lambda\epsilon^3\epsilon^2=f\tau.$$

15. Proof of Lemma (5.7). Consider the commutative diagram

$$\begin{tikzcd}
& QK \arrow[dl, "rQK"'] \arrow[dr, "QrK"] & \\
RQK & K \arrow[u, "qK"'] \arrow[d, "rK"] & QRK \\
& RK \arrow[ul, "RqK"] \arrow[ur, "qRK"'] &
\end{tikzcd}$$

It follows from Definition (5.1b) that the maps rQK and qRK are homotopy equivalences. Let αK (resp. βK) be a homotopy inverse of rQK (resp. qRK). Then the following diagram is commutative up to homotopy

$$\begin{tikzcd}
& QK \arrow[dr, "QrK"] & \\
RQK \arrow[ur, "\alpha K"] & K \arrow[u, "qK"'] \arrow[d, "rK"] & QRK \arrow[dl, "\beta K"] \\
& RK \arrow[ul, "RqK"] &
\end{tikzcd}$$

i.e., $qK \simeq \alpha K \circ RqK \circ rK$ and $rK \simeq \beta K \circ QrK \circ qK$. Consequently

$$qK \simeq (\alpha K \circ RqK) \circ (\beta K \circ QrK) \circ qK,$$
$$rK \simeq (\beta K \circ QrK) \circ (\alpha K \circ RqK) \circ rK.$$

Application of the homotopy extension theorem (which holds for objects of $\mathcal{S}_E$; see [9]) yields c. s. s. maps $s\colon QK \to QK$, $t\colon RK \to RK$ such that

$$s \simeq (\alpha K \circ RqK) \circ (\beta K \circ QrK), \qquad t \simeq (\beta K \circ QrK) \circ (\alpha K \circ RqK)$$

and

$$s(qK)\sigma = (qK)\sigma, \qquad t(rK)\sigma = (rK)\sigma$$

for every simplex $\sigma \in K$. It then follows from condition (5.1c) that s and t are homotopy equivalences. Thus $\alpha K \circ RqK$ and $\beta K \circ QrK$ are homotopy equivalences and because αK and βK are also homotopy equivalences, so are RqK and QrK.

16. Proof of Lemma (8.4). Let $i_{\mathrm{Ex}\,L}\colon L \to \mathrm{Ex}\,L$ be the identity map and let $\mu L = \beta^{-1} i_{\mathrm{Ex}\,L}$. Consider the diagram

$$\begin{tikzcd}
{|K|} \arrow[ddd, "{|h|}"'] & & & {|\mathrm{Sd}\,K|} \arrow[lll, "{|dK|}"'] \arrow[ddll, "{|\mathrm{Sd}\,h|}"] \arrow[ddd, "{|\beta^{-1}h|}"] \\
& & & \\
& {|\mathrm{Sd}\,\mathrm{Ex}\,L|} \arrow[dl, "{|d\mathrm{Ex}\,L|}"] \arrow[drr, "{|\mu L|}"'] & {|\mathrm{Sd}\,L|} \arrow[l, "{|\mathrm{Sd}\,eL|}"'] \arrow[dr, "{|dL|}"] & \\
{|\mathrm{Ex}\,L|} & & & {|L|} \arrow[lll, "{|eL|}"]
\end{tikzcd}$$

In view of the naturality of d commutativity holds in the upper left triangle and the trapezium and because of the naturality of β and the fact that (Lemma (7.2)) $dL = \beta^{-1}(eL)$, commutativity also holds in both triangles which have $|\mu L|$ as lower edge. It follows from Lemma (7.4) and (7.5) that the maps $|dL|$, $|eL|$ and $|d\,\mathrm{Ex}\,L|$ are homotopy equivalences. The commutativity in the trapezium and the smallest triangle involving $|\mu L|$

therefore implies that the maps $|\,Sd\,eL\,|$ and $|\,\mu L\,|$ are also homotopy equivalences. Consequently the lower triangle is commutative up to homotopy and

$$|\,h\,|\circ|\,dK\,| = |\,d\mathrm{Ex}\,L\,|\circ|\,\mathrm{Sd}\,h\,| \simeq |\,eL\,|\circ|\,\mu L\,|\circ|\,\mathrm{Sd}\,h\,| \simeq |\,\epsilon L\,|\circ|\,\beta^{-1}h\,|.$$

THE WEIZMANN INSTITUTE OF SCIENCE, REHOVOTH, ISRAEL
COLUMBIA UNIVERSITY, NEW YORK, N. Y.

REFERENCES.

[1] S. Eilenberg and S. MacLane, "Acyclic models," *American Journal of Mathematics*, vol. 75 (1953), pp. 189-199.

[2] S. Eilenberg and J. A. Zilber, "Semi-simplicial complexes and singular homology," *Annals of Mathematics*, vol. 51 (1950), pp. 499-513.

[3] D. M. Kan, "Abstract homotopy I," *Proceedings of the National Academy of Sciences*, vol. 41 (1955), pp. 1092-1096.

[4] ———, "Abstract homotopy II," *ibid.*, vol. 42 (1956), pp. 225-228.

[5] ———, "Abstract homotopy III," *ibid.*, vol. 42 (1956), pp. 419-421.

[6] ———, "Adjoint functors," *Transactions of the American Mathematical Society*, (to appear).

[7] ———, "Functors involving c. s. s. complexes," *Transactions of the American Mathematical Society*, (to appear).

[8] J. W. Milnor, "The geometric realization of a semi-simplicial complex," *Annals of Mathematics*, vol. 65 (1957), pp. 357-362.

[9] J. C. Moore, "Semi-simplicial complexes and Postnikov systems," *Proceedings International Symposium on Algebraic Topology and its Applications* (1956), Mexico.

[10] J. H. C. Whitehead, "Combinatorial homotopy I," *Bulletin of the American Mathematical Society*, vol. 55 (1949), pp. 213-245.

TOPOLOGICAL ABELIAN GROUPS.*[1]

By Fred B. Wright.

1. **Introduction.** The principal purpose of this note is to establish the existence, in any topological abelian group, of a closed subgroup which has special importance for the structure theory of the group. This subgroup possesses characteristics which are analogous to those of the radical of a ring. It is perhaps appropriate to call this subgroup the radical of the group.

The radical is a well-known subgroup in certain cases. For example, in a discrete abelian group the radical is the torsion subgroup. More generally, in any locally compact abelian group the radical is the union of all compact subgroups. For the additive group of a real linear topological space E, the radical is the set of all elements of E which are annihilated by any continuous linear functional on E.

It is possible to employ the concept of the radical to furnish novel proofs of some theorems which are usually proved by quite different methods. In particular, it is interesting to observe just how much of the structure theory for locally compact abelian groups can be obtained by such methods. In the final section of this paper, it is shown that almost all of this theory can be established intrinsically. That is to say, most of the structure theory can be developed without the use of such "external" constructions as the character group, the L_1-algebra, or the L_2-space of the group.

2. **Preliminary considerations.** Since this paper will be concerned only with abelian groups, it is convenient to use an additive notation for the group operation. In a topological group, it will always be assumed that addition is jointly continuous, that inversion is continuous, and that the Hausdorff separation axiom is satisfied. With some care, the preliminary theory can be developed under weaker restrictions. (A subsequent note will deal with an extension to non-abelian groups.)

This is a convenient place to establish some conventions of notation which will be observed throughout this paper. Let G be a topological abelian group,

---

* Received February 27, 1956; revised January 7, 1957.

[1] Written with the support of the National Foundation grant to Tulane University. The advice and criticism of Professors Goto, Mostert, Pettis and Wallace have contributed much to the present form of this paper.

and let A be any subset of G. Then $\bar{A}$ denotes the closure of A, A^0 denotes the interior of A, and $-A$ denotes the set $\{-a \colon a \in A\}$. A set A is said to be symmetric if $A = -A$. If k is any positive integer, the set kA is defined to be the set $\{a_1 + \cdots + a_k \colon a_i \in A\}$, and $(-k)A = k(-A)$. The complement of A is denoted by $\sim A$, so that $\sim A = \{x \in G \colon x \notin A\}$. The symbol $\square$ denotes the empty subset of G.

If B is another subset of G, then $A \cup B$ and $A \cap B$ denote, respectively, the set-theoretic union and intersection of A and B. The set $\{a + b \colon a \in A, b \in B\}$ is denoted by $A + B$. The set A is a neighborhood of $B \neq \square$ if $B \subset A^0$.

For $x \in G$, the symbol x denotes ambiguously the element x and the set consisting of the single element x. In particular, 0 denotes either the identity element of G or the subgroup consisting of this element only.

A topological isomorphism between two topological groups is an isomorphism between the groups which is also a homeomorphism. Two groups G, H are said to be topologically isomorphic if there is a topological isomorphism between them, and this is denoted by $G \cong H$. The group G is said to be continuously isomorphic to the group H if there is an isomorphism of G onto H which is continuous.

If G and H are two topological abelian groups, their direct sum $G \oplus H$ is the algebraic direct sum endowed with the cartesian product topology.

DEFINITION 2.1. *If A is any nonvoid subset of a topological abelian group G, let $s(A) = \{x \in G \colon x + A \subset A\}$.*

Since $s(A)$ always contains 0, $s(A)$ is nonvoid.

DEFINITION 2.2. *A nonvoid subset S of a topological abelian group G is said to be a semigroup in G if $S \subset s(S)$.*

It is clear that S is a semigroup if and only if $2S \subset S$, which is in accord with the usual definition. It is obvious that for any nonvoid set A the set $s(A)$ is a semigroup in G.

DEFINITION 2.3. *A subset A of a topological abelian group G is said to be a regular open set if A is the interior of its closure; that is, if $A = (\bar{A})^0$. A subset A of G is said to be an angular[2] subset of G if A is open and if $0 \in \bar{A}$.*

If S is a semigroup in G, then $\bar{S}$ is also a semigroup in G, and if $S^0 \neq \square$, then S^0 is also a semigroup in G. If a family of semigroups has

[2] This terminology is due to Hille [4: Definition 7.6.1].

a nonvoid intersection, that intersection is a semigroup. The union of a chain of semigroups is a semigroup. A subset A of G is open (closed, regular, angular) if and only if $-A$ is open (closed, regular, angular); A is a semigroup if and only if $-A$ is a semigroup.

A necessary and sufficient condition that a subset A of G be a subgroup of G is that A be a symmetric semigroup. Another important class of semigroups are those which are completely asymmetric.

DEFINITION 2.4. *A semigroup S in a topological abelian group G is said to be 0-proper if S does not contain the identity element 0 of G.*

Clearly, S is a 0-proper semigroup in G if and only if $S \cap -S = \square$. If S is not 0-proper, then the set $H = S \cap -S$ is a subgroup of G, and is obviously the largest subgroup of G which is contained in S. Then S is a union of cosets of H. If 0 is an interior point of the semigroup S, then $H = S \cap -S$ is an open subgroup of G, so that S is an open set. It follows that in a connected abelian group G an open semigroup is a 0-proper semigroup in G if and only if it is a proper subset of G.

Semigroups of the form $s(A)$, where A is a nonvoid subset of G, have useful properties. It will be observed that $s(-A) = -s(A)$. Then the set $s(A) \cap s(-A)$ is a symmetric semigroup, and hence a subgroup, in G.

DEFINITION 2.5. *For any nonvoid subset A of G, define $b(A)$ to be the subset $s(A) \cap s(-A)$.*

LEMMA 2.1. *For any nonvoid subset A of G, if $x \in b(A)$ then $x + A = A$.*

Proof. If $x \in b(A)$, then $x + A \subset A$ and $-x + A \subset A$. Thus $x + A \subset A \subset x + A$, and hence $A = x + A$.

THOREM 2.2. *If A is a nonvoid regular open set, then $s(\bar{A}) = s(A)$, $s(A)$ is closed, and therefore $b(A)$ is a closed subgroup of G.*

Proof. By definition, $s(\bar{A}) = \cap \{(\bar{A} - x) : x \in \bar{A}\}$, and hence $s(\bar{A})$ is closed. Furthermore, $s(A) \subset s(\bar{A})$, by the continuity of addition. On the other hand, if $x \in s(\bar{A})$, then $x + \bar{A} \subset \bar{A}$, so that $(x + \bar{A})^0 \subset (\bar{A})^0$. Since A is a regular open set, it follows that $x + A \subset A$. Thus $s(\bar{A}) \subset s(A)$, and the proof is complete.

COROLLARY.[3] *If S is a regular open semigroup in G, then $\bar{S} + S \subset S$.*

[3] Compare with [4: Theorem 7.6.3].

LEMMA 2.3. *If S is an angular semigroup in G, then S is a regular open set.*

This is proved in [4: Theorem 7.6.2]. The existence of angular semigroups is sometimes implied by other hypotheses. For example, see [4: Theorem 7.7.1.]

The results of this section are valid for non-commutative groups, provided $s(A)$ is defined to be $\{x \in G : (x + A) \cup (A + x) \subset A\}$.

3. Maximal 0-proper open semigroups. Suppose G is a topological abelian group which contains 0-proper open semigroups. Then Zorn's lemma furnishes semigroups which are maximal with respect to these properties. This section is devoted to the properties of maximal 0-proper open semigroups. The problem of existence is deferred to later sections, and the presence or absence of such semigroups has a direct bearing on the structure of a group.

It will be observed that M is a maximal 0-proper open semigroup in G if and only if $-M$ is such a semigroup.

THEOREM 3.1. *If M is a maximal 0-proper open semigroup in a topological abelian group, then M is a regular open set.*

Proof. If $0 \in \bar{M}$, then M is angular, and hence regular, by Lemma 2.3. Otherwise $\bar{M}$ is a 0-proper semigroup, and hence $(\bar{M})^0$ is an open 0-proper semigroup containing M. Since M is maximal, $M = (\bar{M})^0$.

LEMMA 3.2. *Let M be a maximal 0-proper open semigroup in a topological abelian group G, and let n be any positive integer. If x is such that $nx \in M$, then $x \in M$.*

Proof. Define a function f on G by $f(x) = nx$; then f is continuous. If $N = f^{-1}M$, then N is obviously a 0-proper open semigroup containing M. Hence $M = f^{-1}M$.

THEOREM 3.3. *If M is a maximal 0-proper open semigroup in G, then the set $\sim(M \cup -M)$ is identical with the closed subgroup $b(M)$.*

Proof. Set $B = \sim(M \cup -M)$. Since $s(M) \cap -M = \square$, then $s(M) \subset B \cup M$. Since M is a semigroup, then $M \subset s(M)$. If $x \in B$, then the set $W = \bigcup_{k=0}^{\infty} (kx + M)$ is an open semigroup containing M. If $0 \in W$, then there is a positive integer n such that $-nx \in M$, and by Lemma 3.2, it follows that $-x \in M$. This contradicts the choice of $x \in B$. Thus $0 \notin W$,

and hence $W = M$. In particular, $x + M \subset M$, so that $x \in s(M)$. Thus $B \subset s(M)$, and therefore $M \cup B = s(M)$. Applying this result to $-M$, it follows that $-M \cup B = s(-M)$. Thus $B = s(M) \cap s(-M)$.

This result shows that G can be written as the disjoint union $G = M \cup b(M) \cup -M$, where M and $-M$ are maximal 0-proper open semigroups in G and $b(M)$ is a closed subgroup of G. This fact plays a vital role in all that follows.

The following three results are useful, and are immediate. Their proofs offer no difficulty.

THEOREM 3.4. *If M is a maximal 0-proper open semigroup in G, then $G = M - M$; that is, G is generated by M.*

The last two theorems in this section complete the circle of ideas centering around the notions of maximality, angularity, and regularity of semigroups, and should be compared with Theorem 3.1.

THEOREM 3.5. *Let M be a maximal 0-proper open semigroup in G. Then the following three statements are equivalent:* (1) *M is not angular,* (2) *M is closed,* (3) *$b(M)$ is open.*

THEOREM 3.6. *Let M be a maximal 0-proper open semigroup in G. Then the following are equivalent statements:* (1) *M is angular,* (2) $\bar{M} = M \cup b(M)$, (3) *$b(M)$ is the topological boundary of M.*

The repeated use of subgroups of the form $b(M)$ in the sequel makes it advisable to have a convenient name for them. Theorem 3.3 suggests the following.

DEFINITION 3.1. *A closed subgroup B of a topological abelian group G will be called a residual subgroup of G if there exists a maximal 0-proper open semigroup M of G such that $B = b(M)$; in this case, B is called the residual subgroup belonging to M.*

Every residual subgroup belongs to at least two maximal 0-proper open semigroups in G, and may belong to more.

4. The radical of a topological abelian group. It is now time to introduce the principal definition of this paper.

DEFINITION 4.1. *Let G be any topological abelian group. Denote by T the intersection of all the residual subgroups of G, with the convention that $T = G$ if there are no residual subgroups. The closed subgroup T is*

called the radical of G. If $G = T$, G is said to be a radical group, while if $T = 0$, G is said to be a radical-free group.[4]

THEOREM 4.1. *If G is a discrete abelian group, then the radical T of G is the torsion subgroup of G.*

Proof. For each $x \in G$, the set $\{kx\colon k \geqq 1\}$ is an open semigroup in G. By definition, $x \in T$ if and only if every open semigroup containing x also contains 0. Thus $x \in T$ if and only if x has finite order.

Clearly the radical will always contain the torsion subgroup. It need not, however, be the closure of the torsion subgroup.

A topological abelian group G is a radical group if and only if there are no 0-proper open semigroups. It follows at once that if G is a radical group and if H is a closed subgroup of G, then G/H is a radical group. Thus, if G is a radical group and if S is an open semigroup in G, then $S \cap -S$ is an open subgroup H of G (Section 2). Then G/H is a discrete radical group, and by Theorem 4.1, is a torsion group. Then the image of S in G/H is in fact a group, and, again by Section 2, S is itself a group. This establishes

THEOREM 4.2. *A topological abelian group G is a radical group if and only if every open semigroup in G is an open subgroup of G.*

A non-discrete example of a radical group is therefore furnished by any compact abelian group [15]. Since there are compact groups with no non-zero elements of finite order, this shows that the radical need not be the closure of the torsion subgroup.

If G is a topological abelian group, any subgroup H of G is a topological group in its induced (subspace) topology. Hence H has its own radical, which may be denoted by T_H. It is obvious that $T_H \subset T \cap H$, where T is the radical of G. Whether or not equality holds in general is an open question, except in special cases. It will be seen below that equality of these two groups does hold for locally compact abelian groups.

DEFINITION 4.2. *Let G be a topological abelian group, and let H be any subgroup of G. Then H is said to be a radical subgroup of G if H is a radical group when H is considered to be a topological group in its induced topology.*

Then H is a radical subgroup of G if and only if $H = T_H$. In such a case, $H \subset T$.

[4] Although the terminology "radical" should cause no confusion, it seems advisable to use "radical-free" instead of "semisimple."

LEMMA 4.3. *Let G be a topological abelian group.* (1) *If H is a radical subgroup of G, then $\bar{H}$ is a radical subgroup of G.* (2) *The set-theoretic union of a chain of radical subgroups is a radical subgroup of G.* (3) *The subgroup generated by two radical subgroups of G is also a radical subgroup of G.*

Proof. The proof of (3), for example, will illustrate the proofs of the others. Let H_1 and H_2 be two radical subgroups of G, and let $H = H_1 + H_2$. If M is a maximal 0-proper (relatively) open semigroup in H, then $M \cap H_1 = M \cap H_2 = \square$, and hence both H_1 and H_2 are contained in $b_H(M)$, the residual subgroup in H belonging to M. Thus $H \subset b_H(M)$, and hence no such M can exist.

This yields at once the following result.

THEOREM 4.4. *If G is any topological abelian group, then there exists in G a unique maximal radical subgroup H which is closed and is contained in the radical T of G.*

The question naturally arises: is the radical itself a radical subgroup? In general, this is an open question. For locally compact groups, the answer is affirmative. Another affirmative answer is provided by the next theorem.

THEOREM 4.5. *Let G be a topological abelian group with radical T. If $T \neq G$, the following are equivalent.* (1) *There exists a residual subgroup of G which is open.* (2) *Every residual subgroup of G is open.* (3) *The radical T of G is open. If any one of these is true then the radical T is the unique residual subgroup of G, and T is a radical subgroup of G.*

Proof. It is clear that (3) implies (2) and that (2) implies (1). Finally, assume the residual subgroup $b(M)$ is open, and let S be a 0-proper open semigroup in $b(M)$. Since $b(M) + M = M$, it follows that $S \cup M$ is a 0-proper open semigroup in G, violating the maximality of M. Hence $b(M)$ is a radical group, and hence $b(M) \subset T$, by Theorem 4.4. Since $T \subset b(M)$ in every case, then $T = b(M)$ in this case. This also establishes the last statement of the theorem.

Direct calculation establishes the following behavior of 0-proper open semigroups under certain homomorphisms.

LEMMA 4.6. *Let G be a topological abelian group with radical T, and let H be a closed subgroup of G, such that $H \subset T$. Let G^* denote the quotient group G/H, and let π denote the natural homomorphism of G onto G^*. If M and M^* are 0-proper open semigroups in G and G^* respec-*

tively, then πM *and* $\pi^{-1}M^*$ *are* 0*-proper open semigroups in* G^* *and* G *respectively. If* M *and* M^* *are maximal, then* πM *and* $\pi^{-1}M^*$ *are maximal. In particular, if* M *is a maximal* 0*-proper open semigroup in* G*, then* $M = \pi^{-1}\pi M$.

An immediate corollary of this result is the following useful fact.

THEOREM 4.7. *If* G *is a topological abelian group with radical* T*, and if* H *is a closed subgroup of* G *such that* $H \subset T$*, then the radical of* G/H *is* T/H*. In particular,* G/T *is radical-free.*

The next result characterizes the radical as the minimal subgroup of G satisfying the last statement of the preceding theorem.

THEOREM 4.8. *Let* G *be a topological abelian group with radical* T*. Let* H *be any closed subgroup of* G *such that* G/H *is radical-free. Then* $T \subset H$.

Proof. Let $G^* = G/H$, and let $\pi\colon G \to G^*$ be the canonical homomorphism. If M^* is a maximal 0-proper open semigroup in G^*, then there is a maximal 0-proper open semigroup M in G such that $\pi^{-1}M^* \subset M$. Then $\pi^{-1}b(M^*) \supset b(M)$, where the calculations of residual subgroups are made in the proper groups. Then $\bigcap \pi b(M) \subset \bigcap b(M^*)$, where the right-hand intersection is taken over *all* residual subgroups in G^* and the left-hand intersection is taken over *some* residual subgroups in G. The same inclusion is valid if the left-hand intersection is taken over all residual subgroups of G, obviously. Hence

$$T = \bigcap b(M) \subset \bigcap \pi^{-1}\pi b(M) = \pi^{-1} \bigcap \pi b(M) \subset \pi^{-1} \bigcap b(M^*) = \pi^{-1}0 = H,$$

since G^* is radical-free.

The following manifestly true remarks will be employed from time to time. (1) Any subgroup of a radical-free group is radical-free. (2) A radical-free group remains radical-free if the topology is strengthened. (3) A radical group remains a radical group if the topology is weakened. The last two statements are most useful when translated into terms of continous isomorphisms, and (3) is in fact a special case of a remark following Theorem 4.1 above.

5. Maximally radical-free groups. Certain radical-free groups are of special interest because of their significance in structure theory.

DEFINITION 5.1. *A topological abelian group G is said to be maximally radical-free if the subgroup 0 is a residual subgroup of G.*

Thus G is maximally radical-free if and only if there exists a maximal 0-proper open semigroup M in G with $b(M) = 0$. The class of such groups coincides precisely with the class of linearly ordered abelian groups. For, let A be an abelian group (no topology!) which is linearly ordered by the relation $\geqq$. Let P denote the set of all elements x in A satisfying $x > 0$. Then P is a 0-proper semigroup in A, and every element of A different from 0 is either in P or in $-P$. If, conversely, P is a 0-proper semigroup in G such that $\sim(P \cup -P) = 0$, then there is a linear order in A such that $P = \{x \in A : x > 0\}$. The set P is called the positive cone of A.

A linearly ordered abelian group A has an intrinsic topology defined on it by the order. A base for the neighborhoods of 0 is given by the family of intervals $(-a, a) = \{x \in A : -a < x < a\}$, for all $a \in P$. This topology will be called the interval topology in A. It is easily seen that A is a topological abelian group in the interval topology, and that the positive cone P of A is open.

The interval topology is thus essentially determined by the positive cone P. One minor relation between the algebraic structure of P and the topology in A is worth recording. It is easily seen that either $P = 2P$ or else $P \cap \sim(2P)$ consists of a single element. In the latter case, the group A is discrete in the interval topology, while if $P = 2P$ the group A is dense-in-itself.

Further details about linearly ordered abelian groups may be found in [1]. The remarks above form the skeleton of the very important result contained in the following theorem.

THEOREM 5.1. *Let G be a maximally radical-free topological abelian group. Then G can be made a linearly ordered abelian group, and the interval topology on G is coarser than the original topology. Conversely, if A is a linearly ordered abelian group, then A is a maximally radical-free group in its interval topology.*

This yields an interesting new proof of a classic theorem concerning abelian groups: Any abelian group, all of whose elements have infinite order, can be linearly ordered to be a linearly ordered group [8]. For, if the group is given the discrete topology, it is therefore radical-free (Theorem 4.1). Since the radical is open, it is the unique residual subgroup in the group (Theorem 4.5). Hence the group is maximally radical-free (Definition 5.1).

Thus the group can be made a linearly ordered group (Theorem 5.1). It is interesting to uncover the almost completely hidden use of the axiom of choice, which is usually cited explicitly in proofs of this theorem.

THEOREM 5.2. *A connected maximally radical-free topological abelian group is continuously isomorphic to the additive group of real numbers. If the group is either locally compact or locally connected, then the isomorphism is also a homeomorphism.*

Proof. Such a group has the property that the space obtained by deleting the element 0 is disconnected, by virtue of Theorem 3.3. The conclusion of the theorem follows from [2: page 16, no. 4]. (See also [5] for a precise statement of the theorem cited.)

COROLLARY 1. *Let G be a maximally radical-free group, and let K be the component of 0 in G. Then either $K=0$ or K is continuously isomorphic to the additive group of real numbers.*

In the structure theorems which follow, it will be necessary to establish the fact that a continuous isomorphism is in fact a topological isomorphism. This will be done by citing an appropriate one of the following three references: [14: page 20], [14: page 95], [10: Theorem 6]. The first such structure theorem, which follows from the second of these references, is now at hand.

COROLLARY 2. *Let G be a maximally radical-free group, with K the component of 0 in G. If G/K is discrete, then G is topologically isomorphic to $K \oplus (G/K)$.*

Now let G be maximally radical-free and let K be the component of 0 in G. Then K has the property that for any $x \in K$, with $x > 0$, if $0 < y < x$ then $y \in K$; that is, K is an "isolated subgroup" [12] or a "convex subgroup" [1] of G. To see this, let H be the set of all elements $y \in G$ such that $-x < y < x$ for all $x > 0$ in K. It is easily seen that $H+K$ is the smallest convex subgroup containing K. But if $y \in M$ is different from 0, the interval $(-y, y)$ is an open set in G containing 0. But $(-y,y) \cap K$ is a neighborhod of 0 in K, and $(-y,y) \cap K = 0$. Hence either $K=0$ or $H=0$. In either case K is a convex subgroup, as desired.

Since this is true, the quotient group G/K has an induced linear order [12: Chapter 1, Lemma 2]. Then G/K is maximally radical-free in the interval topology given by this order. It is easy to see that the quotient

topology on G/K, where G is considered in its original topology, is finer than the interval topology on G/K. This yields the following result.

THEOREM 5.3. *If G is a maximally radical-free topological abelian group, and if K is the component of 0 in G, then either $G = K$ or else the quotient group G/K is maximally radical-free.*

Maximally radical-free groups occur in structure theory in the form of certain quotient groups.

THEOREM 5.4. *If G is a topological abelian group and if $b(M)$ is a residual subgroup of G, then $G/b(M)$ is maximally radical-free.*

6. Real linear topological spaces. This section is something of a digression from the prinicipal theme of this paper. It is interpolated here for two reasons. In the first place, the additive group of an infinite dimensional real linear topological space is the most accessible example of a topological abelian group which is not locally compact, but which has almost any other pleasant algebraic or topological property. It is therefore of some interest to see what the radical is for such a group. Of perhaps equal interest is the fact that real linear topological spaces are the natural vehicles for representations of certain abelian groups, as is shown in the next section.

LEMMA 6.1. *Let E be a real linear topological space, and let M be a maximal 0-proper open semigroup in E. Then M is convex.*

Proof. Since M is open, it suffices to show that M is midpoint-convex. Since M is a semigroup, this will follow if it is shown that $x \in M$ implies that $\frac{1}{2}x \in M$. But this follows at once from Lemma 3.2.

This result is sufficient to establish the precise description of a maximal 0-proper open semigroup and its residual subgroup; one needs only to appeal to [7: Theorem 8.10] to obtain the answer.

THEOREM 6.2. *If M is a maximal 0-proper open semigroup in a real linear topological space E, then there exists a continuous linear functional f on E such that $M = \{x \in E : f(x) > 0\}$, $-M = \{x \in E : f(x) < 0\}$, and $b(M) = \{x \in E : f(x) = 0\}$.*

COROLLARY 1. *The set of residual subgroups of a real linear topological space E is in one-to-one correspondence with the set of continuous hyperplanes through 0 in E. The radical of E consists of all those elements of E which are annihilated by every continuous linear functional on E.*

COROLLARY 2.[5] *In any real linear topological space E, if S is an open semigroup in E, then either $S = E$ or S is a subset of a half-space of E.*

It should be noted that the proof of Lemma 6.1 does not depend on either the joint continuity of addition or on the Hausdorff separation axiom. Since [7: Theorem 8.10] also makes no use of these restrictions, this entire section may be extended to such spaces. Theorem 6.2, in particular, shows that the residual subgroups are closed linear subapces. Thus the concept of the radical is applicable to this wider class of linear spaces, and Corollary 1 asserts that a real linear topological space is radical-free if and only if it is a locally convex Hausdorff linear topological space.

7. Representation of connected groups. Throughout this section, let G denote a connected abelian group. If $b(M)$ is a residual subgroup of G, then $G/b(M)$ is continuously isomorphic to the additive group of real numbers. Conversely, if ϕ is a continuous homomorphism of G onto the reals, it is obvious that the kernel $\phi^{-1}0$ of ϕ is the residual subgroup belonging to the pre-image of the open right-half line under ϕ. An easily established necessary and sufficient condition that two continuous homomorphisms ϕ_1 and ϕ_2 of G onto the real line have the same kernel is that there exist a real number α such that $\phi_1(x) = \alpha \cdot \phi_2(x)$ for all $x \in G$.

If E denotes the set of all continuous homomorphisms of G into the real line, then E is a vector space over the reals under pointwise operations. Furthermore, E can be endowed with a topology making it a linear topological space. The most natural, though by no means the only, such topology is the weak topology induced by G. Since G is assumed to be Hausdorff, the space E will also be a Hausdorff space, and it is clearly locally convex in this topology.

Now let E^* denote the dual space of E, and let E^* be given its weak* topology; that is, the topology induced on E^* by E. For each $x \in G$, define $Rx \in E^*$ by setting $Rx(\phi) = \phi(x)$, for each $\phi \in E$. It is clear that R is a homomorphism of G into E^*. Furthermore, $Rx = 0$ if and only if $\phi(x) = 0$ for all $\phi \in E$. As shown above, this holds if and only if x belongs to the radical T of G. The continuity of R follows at once from the definition of the topology in E^*.

Finally, if $\phi \in E$ is such that $\phi(x) = 0$ for all $x \in G$, then $\phi = 0$. This shows that the range of R in E^* generates a linear subspace which is dense in E^* in the weak* topology. These results are summarized as follows.

[5] Compare with [4: Theorem 7.6.8].

THEOREM 7.1. *Let G be a connected abelian group with radical T. Then there exists a locally convex real linear topological space E with dual space E^*, and a continuous homomorphism R of G into E^* such that the kernel of R is T. The range of R generates a weak* dense linear subspace of E^*.*

This result is a generalization of part of the structure theorem for locally compact abelian groups. Furthermore, if G is a normed linear space, the mapping R is the usual injection of G into its second dual space, and the last statement of Theorem 7.1 is Helley's theorem.

If the hypothesis of connectivity is replaced by an algebraic condition, analogous results can be established. The algebraic condition is a generalization of Archimedean order. This topic will be dealt with in a subsequent note.

8. Locally compact abelian groups. This final section is concerned with the application of the methods and results of the preceding sections to the study of locally compact abelian groups. The structure of such groups is well-known [11, 13, 14]. The purpose of this section is to establish as much of this structure theory as possible by means of topological and algebraic methods alone, and within the group itself. This means that the proofs to follow will make no use of measure-theoretic concepts, and that no external constructions will be used. In particular, the character group will not be used. (In the proofs of Theorems 8.3 and 8.4, the concept of a covering group [3] is used. This violates the promise not to use non-intrinsic objects. The fact is that these two results can be given intrinsic proofs which amount to proving the one fact about covering groups which is needed. It has seemed simpler to make this concession to the demands of brevity.)

It must be noted that the full structure theorem has not as yet yielded to these methods. Theorem 8.5, Theorem 8.7 and the Corollary to Theorem 8.9 constitute that part which can be so proved. The missing result is precisely this: the radical of a locally compact connected abelian group is a topological and algebraic direct summand of the group. (One can easily reduce the problem to the case in which the radical is divisible, in which case the radical is algebraically a direct summand.)

To begin, it will be recalled that a compact group is a radical group [15]. It follows that every compact subgroup of a locally compact abelian group is contained in the radical. Hence a radical-free group contains no compact subgroups other than 0.

LEMMA 8.1. *Let G be a totally disconnected radical-free locally compact abelian group. Then G is discrete.*

Proof. Since G is totally disconnected, there exists a compact open subgroup H. Since G is radical-free, $H = 0$, so that G is discrete.

LEMMA 8.2. *Let G be a radical-free locally compact abelian group. Then there exists a finite set of residual subgroups $b(M_1)\cdot\cdot\cdot, b(M_n)$ of G such that the subgroup $B = \bigcap_{j=1}^{n} b(M_j)$ is discrete.*

Proof. Let U be a compact neighborhood of 0, and let W be the boundary of U. When W is compact, and $0 \notin W$. Since G is radical-free, the open sets $(M \cup -M)$, for all maximal 0-proper open semigroups M in G, cover the compact set W. If $(M_1 \cup -M_1), \cdot\cdot\cdot, (M_n \cup -M_n)$ give a covering of W, consider the closed subgroup $B = \bigcap_{j=1}^{n} b(M_j)$. It follows that $B \cap U = B \cap U^0$. If B_K is the component of 0 in B, then $B_K \cap U$ is a nonvoid open and closed subset of B_K. Hence $B_K \cap U = B_K$, and since B_K is radical-free, then $B_K = 0$. Thus B is totally disconnected and therefore, being radical-free itself, it is discrete.

In what follows, the additive group of a real linear topological space is called a vector group, and the linear dimension of the vector space is called the dimension of the vector group.

THEOREM 8.3. *If G is a connected radical-free locally compact abelian group, then G is topologically isomorphic to a vector group of* (*unique*) *finite dimension.*

Proof. Choose the least positive integer such that there exists a collection of n residual subgroups $b(M_1), \cdot\cdot\cdot, b(M_n)$ with discrete intersection B. It is apparent that there is a one-to-one continuous isomorphism of G/B into the direct sum of the groups $G/b(M_i)$, and the minimality of n shows that the mapping is in fact onto this direct sum. Thus G/B is continuously isomorphic to E^n, where E^n is the Euclidean vector group of dimension n. It follows from [9] or [10: Theorem 6] that this mapping is open, and therefore a homeomorphism.

Since G is connected and B is discrete, it follows that G is a covering group for E^n. The uniqueness of the covering group for E^n implies that $B = 0$, and the proof is complete.

THEOREM 8.4. *If G is a connected locally compact abelian group, then the radical T of G is connected.*

Proof. Let H be a subgroup of T which is relatively open in T. By Theorem 4.7 the group G/H has the discrete group T/H as its radical. Thus G/H is a covering group of $(G/H)/(T/H) \cong G/T \cong E^n$. Hence $T/H = 0$, and thus $H = T$.

A complete description of radical-free locally compact abelian groups is now at hand.

THEOREM 8.5. *Let G be a radical-free locally compact abelian group, and let K be the component of 0 in G. Then K is topologically isomorphic to a Euclidean vector group E^n, G/K is discrete, and G is topologically isomorphic to $(G/K) \oplus K$.*

Proof (in outline). Suppose first that G is maximally radical-free. Then G/K is totally disconnected, and is maximally radical-free (Theorem 5.3). Thus, by Lemma 8.1, G/K is discrete, and therefore $G = K \oplus (G/K)$, by Corollary 2 to Theorem 5.2.

Theorem 4.5 shows that consideration may be restricted to the case in which $G/b(M)$ is not discrete for any residual subgroup $b(M)$ of G. Since $G/b(M)$ is maximally radical-free, then $G/b(M) = E^1 \oplus D$, where D is discrete, for each M. Now choose the least integer n such that there exists a family $b(M_i)$, $i = 1, \cdots, n$, of residual subgroups of G having discrete intersection B. Then G/B is topologically isomorphic to $E^n \oplus \Delta_1$, where Δ_1 is discrete. Let $\pi\colon G \to G/B$ be the canonical homomorphism, and let $\Delta = \pi^{-1}\Delta_1$. Then $G/\Delta \cong E^n$, and it is easily seen that $K \cong E^n$. It follows that there is a continuous homomorphism of G onto K leaving K elementwise fixed. By [14: page 20], $G = K \oplus (G/K)$. Then the totally disconnected group G/K must be radical-free (being essentially a subgroup of G), and is therefore discrete.

COROLLARY. *If G is a totally disconnected locally compact abelian group, then the radical T of G is open.*

In particular, in the totally disconnected case the radical is the unique residual subgroup and also the maximal radical subgroup of G. This comment will be extended to the general case later.

The next result is the crucial lemma in the study of radical groups.

LEMMA 8.6. *Let G be a locally compact abelian group, and let K be the component of 0 in G. If K is not compact, then G is not a radical group.*

Proof. Let U be a symmetric open neighborhood of 0 whose closure is compact. Set $U_1 = U$ and define $U_{n+1} = U_n + U$. Then U_n is open and has compact closure. Since K is not compact, neither is G, and therefore $(U_n)^- \neq G$. On the other hand, if $U_n = (U_n)^-$, then U_n is both open and closed, and since it contains 0 it contains K. This, however, implies that K is compact. Hence $U_n \neq (U_n)^-$. The proof can now be completed by copying the proof of Lemma 2, page 153, of [11]. There it is shown that there exists an element d in the boundary of U such that nd is in the boundary of U_n. This being true, it follows that the set $\bigcup_{n=1}^{\infty} (nd + U_n)$ is a 0-proper open semigroup in G.

A more natural way of expressing this result is as follows.

THEOREM 8.7. *If G is a radical locally compact abelian group, then the component of 0 in G is compact.*

COROLLARY. *A necessary and sufficient condition that a connected locally compact abelian group be a radical group is that G be compact.*

It is now necessary to prove directly a piece of the structure theorem. In the usual order of things, this follows as an easy consequence of the theorem. The proof given of this result here is a mild adaptation of the standard construction of a one-parameter subgroup in G.

LEMMA 8.8. *Let G be a locally compact abelian group, and let H be a closed subgroup of G such that H is topologically isomorphic to a vector group E^n, and such that G/H is topologically isomorphic to a vector group E^m. Then G is topologically isomorphic to E^{n+m}.*

Proof. These hypotheses imply that G is connected [3: page 36]. The proof is an induction on m. The case $m = 0$ being trivial, suppose $G/H \cong E^1$. For each integer $k \geqq 1$ and for each $x \in G$, the equation $2^k y = x$ has a unique solution y. (That G is algebraically isomorphic to E^{n+m} follows at once from [6: Theorem 2], for example.) Thus the mappings f_k, defined on G for each positive integer k by $f_k(x) = 2^k x$, are continuous isomorphisms of G onto itself. Since G is connected and locally compact, [10: Theorem 6] applies, and hence f_k is open.

Let U be any compact neighborhood of 0 in G, and let $U_k = f_k^{-1}U$. Then each U_k is a compact neighborhood of 0 in G, and $U_k \supset U_{k+1}$ for each k. Then, given any such U and any $x \in G$ such that $x \notin H$, let $2^{-k}x$ denote the unique solution of the equation $2^k y = x$. If X is the set of all $2^{-k}x$, $k = 1, 2, \cdots$, then $X \subset U$, and since U is compact, $\bar{X}$ is compact and is contained in U. Thus the set X has at least one limit point $a \in U$. It will now be shown that necessarily $a = 0$.

If $2^{-k_i}x \to a$, as $i \to \infty$, it is obvious that for each i, $2^{k_i}a \in U$. One of three cases can occur: (1) $a = 0$, (2) $a \neq 0$ has finite order, (3) U contains an infinite number of non-zero elements in the cyclic subgroup spanned by a. Case (2) is ruled out at once, since G has no such elements. If (3) holds for a in G, it also holds for the image of a in $G/H \cong E^1$. This contradiction implies that $a \in H$, and hence (3) holds in $H = E^n$. This is also impossible. Therefore $2^{-k}x \to 0$ and $k \to \infty$.

A copy of the reals may now be constructed in G. The group generated by X is disjoint from H, and is a continuously isomorphic copy of the dyadic rationals. Then this isomorphism can be extended to a continuous isomorphism of the reals E^1 into G. The isomorphism is necessarily a homeomorphism (onto its image), and it readily follows that $G \cong E^n \oplus E^1 \cong E^{n+1}$. An easy induction establishes the general result.

It is now possible to establish the important fact that the radical of any locally compact abelian group is a radical subgroup.

THEOREM 8.9. *The radical T of a locally compact abelian group G is the maximal radical subgroup of G.*

Proof. If T is not a radical group, then there is a closed subgroup $H \subset T$ such that H is a residual subgroup of T. Then G/H has T/H as its radical. This reduces the problem to showing that the radical cannot be maximally radical-free. A similar reduction enables one to consider separately the possibility that T is either the additive group of reals or is a discrete torsion-free group.

Suppose first that $T \cong E^1$. Then $T \subset K$, where K is the component of 0 in G. Then $G/K \cong (G/T)/(K/T)$. Since K is connected, it follows that $K/T \cong E^n$ for some integer n. By Lemma 8.8, $K \cong E^{n+1}$. But $G/T \cong E^m \oplus \Delta$, for integer m and a discrete group Δ. Then $G/K \cong E^{m-n} \oplus \Delta$. Since G/K is totally disconnected, $m - n = 0$, so that G/K is discrete. Since K is divisible, $G \cong K \oplus (G/K)$. But then G has radical 0, contrary to assumption.

This leaves the possibility that the radical T is a discrete torsion-free group. As before, $G/T \cong E^m \oplus \Delta$. If K is the component of 0, then $K \cap T$ is a closed discrete subgroup of K containing the radical of K. Since the radical of K is connected, K is radical-free, and hence $K \cong E^n$ for some integer $n \geqq 0$. On the other hand, the group $K/(K \cap T)$ is mapped continuously and isomorphically into the component E^m of G/T. Thus $K/(K \cap T)$ contains no elements of finite order. But if $K \cap T \neq 0$, $K/(K \cap T)$ contains a non-trivial torus, and thus has elements of finite order. Hence $K \cap T = 0$.

Let $H = (K + T)^-$. Since $H/T \subset G/T \cong E^m \oplus \Delta$, then $H/T \cong E^k \oplus \Delta_1$, where Δ_1 is discrete and $k \leqq m$. In the projection of H onto H/T, K is mapped continuously and isomorphically into the component E^k of H/T. Since $G/H \cong (G/T)/(H/T)$, it follows that $G/H \cong E^r \oplus Z^s \oplus \Delta_2$, where Z^s denotes a torus of dimension s, where Δ_2 is discrete, and where $r + s = m - k$. But the left-hand side is totally disconnected, so that $r + s = 0$, and hence $m = k$. Thus $H/T \cong E^m \quad \Delta_1$. Let π denote the projection of H onto H/T. Since π is continuous and since $H = (K + T)^-$, then $E^m \oplus \Delta_1 \cong H/T = \pi H = \pi(K + T)^- \subset (\pi(K + T))^- = (\pi K)^- \subset E^m$. Thus $\Delta_1 = 0$ and $H/T \cong E^m$. Then π yields a continuous isomorphism of $K \cong E^n$ onto a dense subset of E^n. By virtue of [2: Chapter VII, § 2, no. 1], K is closed, so that $H/T \cong E^n \cong K$. It follows at once from [14: page 20] that $H \cong K \oplus \Delta_3$, where Δ_3 is discrete and topologically isomorphic to H/K.

Since G/H is discrete, H is open in G, and since H/K is discrete, K is open in H. Hence K is open in G, so that G/K is discrete. By [14: page 95], $G \cong K \oplus \Delta_4$, where $\Delta_4 \cong G/K$. Since $\Delta_4$is discrete, its radical is its torsion subgroup,. which is contained in T. But T is torsion-free, and hence G has radical 0, contrary to assumption. This completes the proof.

COROLLARY. *If G is a connected locally compact abelian group, then the radical T of G is a connected compact subgroup having the property that G/T is topologically isomorphic to a vector group of (unique) finite dimension.*

THEOREM 8.10. *Let G be a locally compact abelian group. Then the following sets in G are identical:* (1) *the radical T of G,* (2) *the maximal radical subgroup of G,* (3) *the set of all elements x in G such that the cyclic subgroup of G generated by x has compact closure.*

Proof. The equality of (1) and (2) was established in Theorem 8.8.

Clearly set (3) is contained in set (1). Conversely, let $x \in T$, and let K be the component of 0 in T. Then K is compact, and the cyclic group (x) has compact closure if $x \in K$. If $x \notin K$, let x^* be the image of x in the radical group $T^* = T/K$. If $\pi\colon T \to T^*$ is the canonical projection, then $\pi(x)^- \subset (\pi x)^- = (x^*)^-$, so that $(x)^- \subset \pi^{-1}(x^*)^-$. If $(x^*)^-$ is known to be compact, then $\pi^{-1}(x^*)^-$ is also compact, since K is compact. This reduces the question to the totally disconnected case.

If T is totally disconnected, let $x \in T$ and let H be a compact open subgroup of T. Then T/H is a discrete radical group, and hence a torsion group. Then there is some positive integer k such that $kx \in H$. Thus $(x) \subset \bigcup_{i=1}^{k} (ix + H)$, and hence $(x)^-$ is compact.

COROLLARY. *If G is a locally compact abelian group with radical T, and if H is a closed subgroup of G, then the radical of H is $H \cap T$.*

It should be remarked that a more natural order of theorems would be to have Theorem 8.10 follow immediately after Theorem 8.5. In this order, Theorem 8.9 could be eliminated, and Theorem 8.7 could be established with considerable ease (modifying the argument for the result on page 97 of [14]). A direct proof of Theorem 8.10 would, however, be quite similar to the proof of Lemma 8.8.

TULANE UNIVERSITY.

REFERENCES.

[1] G. Birkhoff, "Lattice-ordered groups," *Annals of Mathematics*, vol. 43 (1942), pp. 298-331.

[2] N. Bourbaki, *Topologie Générale* (Actualités Scientifiques et Industrielles, No. 1029), Paris, 1947.

[3] C. Chevalley, *Theory of Lie Groups. I.* Princeton, 1946.

[4] E. Hille, *Functional Analysis and Semigroups*, American Mathematical Society Colloquium Publications, vol. 31, New York, 1948.

[5] T. Ganea, *Zur Charakterisierung einparametriger topologischer Gruppen*, Com. Acad. R. P. Române 1 (1951), pp. 731-732 (Math. Rev. 17 (1956), p. 60).

[6] I. Kaplansky, *Infinite Abelian Groups*, University of Michigan Publications in Mathematics, No. 2, Ann Arbor, 1954.

[7] V. L. Klee, "Convex sets in linear spaces," *Duke Mathematical Journal*, vol. 18 (1951), pp. 443-466.

[8] F. Levi, "Arithmetische Gesetze im Gebiete diskreter Gruppen," *Rendiconti del Circolo Matematico di Palermo*, vol. 35 (1913), pp. 225-236.

[9] E. J. McShane, "Images of sets satisfying the condition of Baire," *Annals of Mathematics*, vol. 52 (1950), pp. 293-308.

[10] B. J. Pettis, "On continuity and openness of homomorphism in topological groups," *Annals of Mathematics*, vol. 52 (1950), pp. 293-308.

[11] L. Pontrjagin, *Topological Groups*, Princeton, 1946.

[12] O. F. G. Schilling, *The Theory of Valuations*, Mathematical Surveys, No. 4, New York, 1950.

[13] E. R. van Kampen, "Locally bicompact abelian groups and their character groups," *Annals of Mathematics*, vol. 36 (1935), pp. 448-463.

[14] A. Weil, *L'intégration dans les groupes topologiques et ses applications*, Paris, 1940.

[15] F. B. Wright, "Semigroups in compact groups," *Proceedings of the American Mathematical Society*, vol. 7 (1956), pp. 309-311.

ON CURVES OF MINIMAL LENGTH WITH A CONSTRAINT ON AVERAGE CURVATURE, AND WITH PRESCRIBED INITIAL AND TERMINAL POSITIONS AND TANGENTS.*[1]

By L. E. Dubins.

1. Introduction and summary. Let a particle pursue a continuously differentiable path from an initial point u to a terminal point v. Suppose that its speed is unity and suppose that its velocity vectors at u and v are U and V respectively. We are interested in a path of minimal length for the particle. It is easy to see that there exist u, U, v and V for which no path of minimal length exists. We need some further reasonable restriction. At first, it seems natural to require that the path possess a curvature everywhere, and to prescribe that its radius of curvature be everywhere greater than or equal to a fixed number R. But again there exist (u, U, v, V, R) for which no path of minimal length exists (Proposition 14). The difficulty is that we have imposed too severe a restriction. In order to arrive at the correct restriction to impose, we observe that if X is a curve in real n-dimensional Euclidean space, parameterized by arc length, for which $X''(s)$ exists everywhere, then the curvature, $\| X''(s) \|$, is less than or equal to R^{-1} everywhere, if and only if,

$$\| X'(s_1) - X'(s_2) \| \leqq R^{-1} \mid s_1 - s_2 \mid, \tag{1}$$

for all s_1 and s_2 in the interval of definition of X. By the *average curvature* of X in the interval $[s_1, s_2]$ we mean the left side of (1) divided by $\mid s_1 - s_2 \mid$. We say that a curve X in real Euclidean n-space parameterized by arc length has *average curvature always less than or equal to* R^{-1} provided that its first derivative X' exists everywhere and satisfies the Lipschitz condition (1). We inquire, for fixed vectors u, U, v, V in real n-dimensional Euclidean space, E_n, and a fixed positive number R, as to the existence and nature of a path of minimal length among the curves in E_n, of average curvature everywhere less than or equal to R^{-1}. Now we find that paths of minimal length necessarily exist. We call such a path an *R-geodesic.* The purpose of this paper

* Received April 28, 1956; revised January 3, 1957.

[1] This research was supported in part by the United States Air Force through the Air Force Office of Scientific Research of the Air Research and Development Command.

is to prove Theorem 1, which implies that for $n=2$, an R-geodesic is necessarily a continuously differentiable curve which consists of not more than three pieces, each of which is either a straight line segment or an arc of a circle of radius R. Furthermore, the corollary to Theorem 1 implies that three is the least integer for which this is true. The nature of R-geodesics for $n \geqq 3$ is open.

2. Existence of *R*-geodesics. Let u, v, U and V be vectors in real n-dimensional Euclidean space, E_n. Let $\| U \| = \| V \| = 1$ and let $R > 0$. Let $C = C(n, u, U, v, V, R)$ be the collection of all curves X defined on a closed interval $[0, L]$, where $L = L(X)$ varies with X, such that $X(s) \in E_n$ for $0 \leqq s \leqq L$; $\| X'(s) \| \equiv 1$; the average curvature of X is everywhere less than or equal to R^{-1}; $X(0) = u$, $X'(0) = U$, $X(L) = v$ and $X'(L) = V$.

PROPOSITION 1. *For any n, u, U, v, V, and R, there exists an X in $C = C(n, u, U, v, V, R)$ of minimal length.*

Proof. We omit the verification that C is non empty. Let $X_1 \in C$. Let d_1 be the length of X_1 and let d be the infimum of lengths of all curves in C. Clearly $d \leqq d_1$. There exists a sequence $X_n \in C$ such that the length d_n of X_n is monotonely decreasing to d. Since $\| X_n'(s) \| \equiv 1$, it follows that the X_n' are a uniformly bounded family of functions. Since $\| X_n'(s_1) - X_n'(s_2) \| \leqq R^{-1} | s_1 - s_2 |$ for all s_1 and s_2 in the interval $[0, d]$, it follows that the X_n' also form an equicontinuous family on $[0, d]$. Therefore by Ascoli's theorem, [1], there is a subsequence of the X_n whose derivatives converge uniformly on $[0, d]$ to a function Y. For convenience, we assume that X_n is itself such a sequence. It is easy to see that $Y(0) = U$ and $\| Y(s_1) - Y(s_2) \| \leqq R^{-1} | s_1 - s_2 |$ for all s_1 and s_2 in $[0, d]$. Since

$$X_n(s) = X_n(0) + \int_0^s X_n'(t)\,dt = u + \int_0^s X_n'(t)\,dt$$

it follows that for $0 \leqq s \leqq d$,

$$\| X_n(s) - X_m(s) \| \leqq \int_0^s \| X_n'(t) - X_m'(t) \| \,dt \leqq d \cdot \sup \| X_n'(t) - X_m'(t) \|,$$

where the sup is taken over t in $[0, d]$. Therefore X_n converges uniformly in $[0, d]$ to a function X. Since X_n' converges uniformly to Y and X_n converges to X it follows that $X' = (\lim X_n)' = \lim X_n' = Y$. It is elementary to complete the proof of the theorem by showing that $X(d) = v$ and $X'(d) = V$. Namely,

$$\| X(d) - v \| \leqq \| X(d) - X_n(d) \| + \| X_n(d) - v \|$$
$$= \| X(d) - X_n(d) \| + \| X_n(d) - X_n(d_n) \| \leqq \| X(d) - X_n(d) \| + d_n - d$$

which converges to zero as n approaches infinity. Also,

$$\| X'(d) - V \| \leqq \| X'(d) - X_n'(d) \| + \| X_n'(d) - V \|$$
$$= \| X'(d) - X_n'(d) \| + \| X_n'(d) - X_n'(d_n) \| \leqq \| X'(d) - X_n'(d) \| + R^{-1} | d_n - d |$$

which also converges to zero as n approaches ∞. Thus $X(d) = v$, $X'(d) = V$, and therefore $X \in C$, and X is of minimal length.

3. Some preliminary propositions. The purpose of this section is to prove Proposition 6. We begin by borrowing ideas from E. Schmidt's proof, [2], of A. Schur's Lemma and thereby prove:

PROPOSITION 2. *Let X be a curve of average curvature everywhere less than or equal to R^{-1} and Let Z be a semicircle of radius R. Then*

$$\| X'(s) - X'(t) \| \leqq \| Z'(s) - Z'(t) \|, \tag{2}$$

and

$$(X'(s), X'(t)) \geqq (Z'(s), Z'(t)) \tag{3}$$

for all s and t in $[0, \pi R]$. Furthermore, for any fixed s and t, $s < t$, equality holds in (2) if and only if equality holds in (3), and equality holds in (3) if and only if $X(r)$ is an arc of a circle of radius R for $s \leqq r \leqq t$.

Proof. Condition (1) implies that $X''(s)$ exists almost everywhere and is a measurable function of s bounded by R^{-1}. Therefore

$$\left| \int_s^t \| X''(r) \| \, dr \right| \leqq R^{-1} | s - t |.$$

Since X' is absolutely continuous, we see that the arc length of the curve $X'(r)$ for $s \leqq r \leqq t$ is less than or equal to $R^{-1} | s - t |$. Sincs $X'(r)$ is a curve on the surface, S_n, of the unit sphere in E_n, it follows that the length of a geodesic on this surface which connects $X'(s)$ and $X'(t)$ is certainly less than or equal to $R^{-1} | s - t |$. That is, the great circle on S_n containing both $X'(s)$ and $X'(t)$ is divided into two arcs by $X'(s)$ and $X'(t)$. The length of the smaller of these is $\leqq R^{-1} | s - t |$. It is well known that if two great circular arcs on the unit sphere in E_n are each of length less than or equal to π, then the length of the chord subtended by the smaller arc is less than or equal to the length of the chord subtended by the larger arc.

Therefore, for $R^{-1}|s-t| \leqq \pi$, $\|X'(s)-X'(t)\|$ is less than or equal to the length of the chord subtended by a great circular arc of length $R^{-1}|t-s|$. It is easy to see that for $|t-s| \leqq \pi R$, $\|Z'(s)-Z'(t)\|$ is the length of such a chord. Therefore $\|X'(s)-X'(t)\| \leqq \|Z'(s)-Z'(t)\|$ for $|s-t| \leqq R\pi$. This completes the proof of the first part of the proposition.

It is trivial that if $X(r)$ is an arc of a circle of radius R for $s \leqq r \leqq t$, then $\|X'(s)-X'(t)\| = \|Z'(s)-Z'(t)\|$. Assume, therefore, for some s and t, $0 \leqq s < t \leqq \pi R$, that $\|X'(s)-X'(t)\| = \|Z'(s)-Z'(t)\|$. It follows that the length of the smaller great circular arc between $X'(s)$ and $X'(t)$ equals $R^{-1}|s-t|$. Therefore the arc length of the curve X', for $s \leqq r \leqq t$, is $\geqq R^{-1}|s-t|$. Since it was previously shown to be $\leqq R^{-1}|s-t|$, we conclude that it equals $R^{-1}|s-t|$. Therefore X' is a geodesic on S_n connecting $X'(s)$ with $X'(t)$. Therefore $X'(r)$ is an arc of a unit circle of length $R^{-1}|s-t|$ for $s \leqq r \leqq t$. It now easily follows that $X(r)$ is an arc of a circle of radius R for $s \leqq r \leqq t$. The remainder of the proof of the proposition is trivial.

PROPOSITION 3. *Let X be a curve of average curvature everywhere less than or equal to R^{-1} and let Z be a semicircle of radius R. Let γ be the vector of length R determined by the condition that $Z(0)+\gamma$ is the center of the semicircle Z. Let λ be any vector of length R orthogonal to $X'(0)$. Then*

$$(1) \qquad (Z'(s),\gamma) \geqq (X'(s),\lambda)$$

for $0 \leqq s \leqq \frac{1}{2}\pi R$. Furthermore equality holds for some s in this interval if and only if $X(r)$ is an arc of a circle of radius R for $0 \leqq r \leqq s$, and λ is the vector of length R determined by the condition that $X(0)+\lambda$ is the center of the circle determined by X.

Proof. It is easy to see that $Z'(s)$ is a linear combination of the two unit orthogonal vectors $Z'(0)$ and γ/R. Therefore

$$(2) \qquad 1 = (Z'(s),Z'(s)) = (Z'(s),Z'(0))^2 + (Z'(s),\gamma/R)^2.$$

By Proposition 2, we have,

$$(3) \qquad (X'(s),X'(0)) \geqq (Z'(s),Z'(0)).$$

Since $(Z'(s),Z'(0)) \geqq 0$ for $0 \leqq s \leqq \frac{1}{2}\pi R$, we see that (3) implies

$$(4) \qquad (X'(s),X'(0))^2 \geqq (Z'(s),Z'(0))^2.$$

Combining (2) and (4) we get

$$(5) \qquad 1 \leqq (X'(s),X'(0))^2 + (Z'(s),\gamma/R)^2.$$

Since λ/R and $X'(0)$ are unit orthogonal vectors, it follows that

$$(6) \qquad 1=(X'(s),X'(s))\geqq(X'(s),X'(0))^2+(X'(s),\lambda/R)^2.$$

From (5) and (6) we immediately obtain

$$(7) \qquad (Z'(s),\gamma/R)^2\geqq(X'(s),\lambda/R)^2.$$

Since $(Z'(s),\gamma/R)\geqq 0$ for $0\leqq s\leqq\frac{1}{2}\pi R$, we conclude from (7) that

$$(8) \qquad (Z'(s),\gamma)\geqq(X'(s),\lambda)$$

for $0\leqq s\leqq\frac{1}{2}\pi R$. This proves the first part of the proposition.

Assume now, for some s, $0\leqq s\leqq\frac{1}{2}\pi R$, that equality holds in (8). It is easy to see that, therefore, equality must hold in (7). This implies that equality must hold in both (5) and (6) for this particular s. From the equality in (5) we obtain equality in (4) and hence, equality in (3). Therefore, by Proposition 2, $X(r)$ is an arc of a circle of radius R for $0\leqq r\leqq s$. Equality in (6) implies that $X'(s)$ is spanned by $X'(0)$ and λ/R. Since λ is orthogonal to $X'(0)$ it is either the vector determined by the condition that $X(0)+\lambda$ is the center of the semicircle X or the negative of this vector. For $0<s\leqq\frac{1}{2}\pi R$, it is not the negative of this vector since $0<(Z'(s),\gamma)=(X'(s),\lambda)$.

Proposition 4. *Let X be a curve with average everywhere less than or equal to R^{-1} and let λ be any vector of length R orthogonal to $X'(0)$. Then*

$$(X(s)-X(0)-\lambda,X'(s))\geqq 0$$

for $0\leqq s\leqq\frac{1}{2}\pi R$. Furthermore, equality holds for some s in this interval if and only if $X(r)$ is an arc of a circle of radius R for $0\leqq r\leqq s$ and λ is the vector of length R determined by the condition that $X(0)+\lambda$ is the center of the circle determined by X.

Proof. It is easy to see that

$$(9) \qquad (X(s)-X(0)-\lambda,X'(s))=\int_0^s(X'(t),X'(s))dt-(\lambda,X'(s)),$$

for

$$(10) \qquad \begin{aligned}(X(s)-X(0)-\lambda,X'(s))&=(X(s)-X(0),X'(s))-(\lambda,X'(s))\\&=(\int_0^s X'(t)dt,X'(s))-(\lambda,X'(s))\\&=\int_0^s(X'(t),X'(s))dt-(\lambda,X'(s)).\end{aligned}$$

Let $Z(s)$ and γ be as in the hypothesis of Proposition 3. Then Propositions 2 and 3 imply that (10) is greater than or equal to

$$\int_0^s (Z'(t), Z'(s))\,dt - (\gamma, Z'(s)). \tag{11}$$

It is easy to see that (11) is equal to

$$(Z(s) - Z(0) - \gamma, Z'(s)) \tag{12}$$

for the argument is the same as the one which established equality in (9). We now observe that (12) equals one half of

$$d(Z(s) - Z(0) - \gamma, Z(s) - Z(0) - \gamma)/ds, \tag{13}$$

which in turn is equal to

$$d\,\| Z(s) - Z(0) - \gamma \|^2/ds = dR^2/ds = 0. \tag{14}$$

We now prove the second part of the theorem. Assume that for some s, $0 < s \leqq \frac{1}{2}\pi R$,

$$(X(s) - X(0) - \lambda, X'(s)) = 0. \tag{15}$$

Thus the inequalities from (9) through (14) become equalities. In particular, (10) equals (11). But in virtue of Propositions 2 and 3, equality in (10) and (11) imply:

$$\int_0^s (X'(t), X'(s))\,dt = \int_0^s (Z'(t), Z'(s))\,dt; \tag{16}$$

and

$$(\lambda, X'(s)) = (\gamma, Z'(s)). \tag{17}$$

Thus if (15) holds, (17) necessarily also holds. We now invoke the second part of Proposition 3 to complete the proof of the non-trivial part of the present theorem.

Our previous propositions are valid for curves X in any Euclidean space. However, our next two propositions deal only with planar curves. At each point of a differentiable planar curve X there are two tangent circles of radius R. The curve X induces on each of these circles an orientation, so that one of these circles is oriented clockwise, the other counterclockwise. *Let Z_s and Y_s be, respectively, the counterclockwise and clockwise oriented circles of radius R, tangent to the curve X at the point $X(s)$.*

PROPOSITION 5. *Let X be a planar curve with average curvature everywhere less than or equal R^{-1}. Let $D(s)$ be the distance between the center*

of the circle Y_s and the center of the circle Z_0. Then $D(s)$ is a monotone non-decreasing function of s for $0 \leqq s \leqq \frac{1}{2}\pi R$. Furthermore, $D(s) = D(0)$ for some s in this interval if and only if $X(r)$ is a continuously differentiable curve in $[0, s]$ such that, for some r_0 in the closed interval $[0, s]$, $X(r)$ is a counterclockwise oriented arc of a circle of radius R for $0 \leqq r \leqq r_0$ and $X(r)$ is a clockwise oriented arc of a circle of radius R for $r_0 \leqq r \leqq s$.

Proof. Let $f(s)$ equal $D^2(s)$. We show that f is non-decreasing. Let T be a rotation through an angle of $\frac{1}{2}\pi$ in the counterclockwise direction. Clearly

$$D(s) = \| X(s) - RT(X'(s)) - X(0) - RT(X'(0)) \|. \tag{18}$$

Since both X and X' are absolutely continuous and the product of two absolutely continuous functions is likewise absolutely continuous, it follows that $f(s)$ is absolutely continuous. Furthermore

$$f'(s) = 2(X(s) - RT(X'(s)) - X(0) - RT(X'(0)), X'(s) - RTX''(s)), \tag{19}$$

Since $T(X'(s))$ is orthogonal to both $X'(s)$ and $T(X''(s))$, it follows that

$$f'(s) = 2(X(s) - X(0) - RTX'(0), X'(s) - RTX''(s)) \tag{20}$$

for all s for which $X''(s)$ exists. Furthermore, $-T(X''(s))$ is some scalar multiple of $X'(s)$, say, $-T(X''(s)) = k(s)X'(s)$. Therefore

$$f'(s) = 2(X(s) - X(0) - RTX'(0), X'(s) + Rk(s)X'(s)) \tag{21}$$

or equivalently,

$$f'(s) = (1 + Rk(s))(X(s) - X(0) - RTX'(0), X'(s)) \tag{22}$$

for all s such that $k(s)$ exists. Furthermore, since $\| X''(s) \| \leqq R^{-1}$,

$$1 + Rk(s) \geqq 0. \tag{23}$$

We easily conclude from (22), (23) and Proposition 4 that $f'(s) \geqq 0$ almost everywhere. Since we already showed that $f(s)$ is absolutely continuous it follows that $f(s)$ is montone non-decreasing. Hence, so is $D(s)$.

This completes the proof of the first part of the theorem. Now assume that $D(s) = D(0)$ for some s with $0 < s \leqq \frac{1}{2}\pi R$. Since D is monotone it follows that $D(r) = D(0)$ for all r with $0 \leqq r \leqq s$. Therefore, $f(r) = D^2(r)$ is constant and consequently $f'(r) = 0$ for $0 \leqq r \leqq s$. It follows from (22) that for every r for which $k(r)$ exists, either

(24) $\quad 1 + Rk(r) = 0$ or (25) $\quad (X(r) - X(0) - RTX'(0), X'(r)) = 0.$

Let r_0 be the least upper bound of the set of r in $[0, s]$ for which (25) holds. By continuity of X and X', it is easy to show that (25) also holds for $r = r_0$. By the second part of Proposition 4, we conclude that $X(r)$ is an arc of a circle of radius R for $0 \leqq r \leqq r_0$. Again by (25), we conclude that this arc is counterclockwise oriented. Suppose $r_0 < s$. Then for all r in the half-closed interval $(r_0, s]$ for which $k(r)$ exists, (24) holds. Thus $k(r) = -R^{-1}$ for all r in $(r_0, s]$ for which $k(r)$ exists. Since X' is absolutely continuous, it is easy to prove, therefore, that $X(r)$ is an arc of a circle of radius R for $(r_0, s]$. Since $k(r)$ is negative this arc is clockwise oriented. Since X and X' are continuous at $r = r_0$, the proposition is proven.

As an immediate corollary to Proposition 5 we obtain the main result of this section:

PROPOSITION 6. *Let X be a planar curve with average curvature everywhere less than or equal to R^{-1} and let s be any point in the closed interval $[0, \frac{1}{2}R\pi]$. Then the circle Y_s is either disjoint from, or tangent to, the circle Z_0. Furthermore Y_s is tangent to Z_0 if and only if $X(r)$ is a continuously differentiable curve for $0 \leqq r \leqq s$ such that, for some r_0 in the closed interval $[0, s]$, $X(r)$ is a counterclockwise oriented arc of a circle of radius R for $0 \leqq r \leqq r_0$, and $X(r)$ is a clockwise oriented arc of a circle of radius R for $r_0 \leqq r \leqq s$.*

It is of course obvious that there is a proposition similar to Proposition 6 which is concerned with the circles Z_s and Y_0, rather than Z_0 and Y_s. We do not state this theorem but will also refer to it as Proposition 6 in the sequel.

4. Certain curves are *R*-geodesics. The purpose of this section is to prove Proposition 9 which states that certain curves, composed of arcs of circles of radius R and straight line segments, are R-geodesics.

Let $X(s)$ be a convex arc defined for $a \leqq s \leqq b$. Let m be the line determined by the two points $X(a)$ and $X(b)$ and let T be the perpendicular projection onto the line m. We will say that a *point p is above the curve X* provided the line segment whose end points are p and $T(p)$ contains a point of the arc $X(s)$. And we will say that a *curve $Y(s)$* defined for $c \leqq s \leqq d$ *lies above the curve $X(s)$* provided that the image under T of all point of Y which lie above the curve X is the segment whose end points are $X(a)$ and $X(b)$.

We state without proof the following proposition which is geometrically obvious.

PROPOSITION 7. *Let X and Y be planar curves defined on the intervals*

$[a, b]$ and $[c, d]$ respectively. Suppose that X is a convex arc and that $X(a) = Y(c)$ and $X(b) = Y(d)$. Then if Y lies above X the arc length of Y is not less than the arc length of X, and equality holds if and only if Y is a one-one parameterization of the range of X.

As an immediate corollary we have:

PROPOSITION 8. (See fig. 1.) *Let d be the length of the smaller arc of a circle Z determined by two points A and B on Z. Let m be the half*

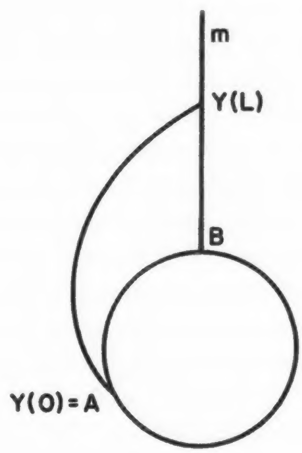

Figure 1.

ray perpendicular to Z at B which does not meet the interior of Z. Suppose $Y(s)$ is a continuous curve parameterized by arc length for $0 \leqq s \leqq L$ which has no point in the interior of Z. Suppose that $Y(0) = A$ and $Y(L)$ is on m. Then $L \geqq d$ and equality holds if and only if $Y(s)$ is a 1-1 parameterization of the smaller arc AB.

We will say that a continuously differentiable curve Y, with arc length s as parameter, defined for $a \leqq s \leqq d$, is of *type ALA* (Arc, Line, Arc) provided there exist b and c with $a \leqq b \leqq c \leqq d$ such that Y restricted to $[a, b]$ is a 1-1 parameterization of an arc of a circle of radius R of length less than or equal to $\frac{1}{2}R\pi$, and such that Y restricted to $[b, c]$ is a line segment, and such that Y restricted to $[c, d]$ is also an arc of a circle of radius R of length less than or equal to $\frac{1}{2}R\pi$.

PROPOSITION 9. *Let u, v, U, and V be vectors in Euclidean 2 space with $\| U \| = \| V \| = 1$ and let R be a positive real number. Suppose Y is*

of type ALA defined for $0 \leqq s \leqq d$, *and suppose* $Y(0) = u$, $Y'(0) = U$, $Y(d) = v$ *and* $Y'(d) = V$. *Then Y is the unique R-geodesic in the collection* $C = C(2, u, U, v, V, R)$.

Proof. Suppose that $0 \leqq d_1 \leqq d_1 + d_2 \leqq d$ and that Y restricted to $[0, d_1]$ is an arc of a circle, Z_1, of radius R and length $\leqq \frac{1}{2}R\pi$, and that Y restricted to $[d_1, d_1 + d_2]$ is a line segment, and Y restricted to $[d_1 + d_2, d]$ is an arc of a circle, Z_2, of radius R and length less than or equal to $\frac{1}{2}R\pi$. Let m_1 be the line passing through $Y(d_1)$ and perpendicular to Y at $Y(d_1)$ and let m_2 be the line parallel to m_1 which passes through $Y(d_1 + d_2)$. The plane is divided by m_1 and m_2 into three strips. Now let X be any curve in C. By a connectedness argument, it is easy to see that there is a least positive number s_1 such that $X(s_1)$ is on m_1. We claim $s_1 \geqq d_1$. For either, for all s, $0 \leqq s \leqq s_1$, $X(s)$ is not in the interior of the circle Z_1, or, for some $s_0 < s_1$, $X(s_0)$ is in the interior of Z_1. In the first case, the preceding proposition implies that $s_1 \geqq d_1$. In the second case, Projosition 6 implies that $s_0 > \frac{1}{2}\pi R$. Since $s_1 > s_0 > \frac{1}{2}\pi R \geqq d_1$, it follows that $s_1 > d_1$. Again by a connectedness argument, there is a smallest number s_2 such that $X(s_2)$ is on m_2. Since m_1 and m_2 are a distance d_2 apart, it follows that $s_2 \geqq s_1 + d_2$. Suppose L is any number such that $X(L) = Y(d)$. An argument similar to the one which showed $s_1 \geqq d_1$ will prove that $L - s_2 \geqq d - (d_1 + d_2)$. Therefore $L \geqq d$. Thus we've shown that Y is an R-geodesic in C. Furthermore the conditions for equality in the preceding proposition imply that Y is the unique R-geodesic in C. This completes the proof.

5. Another preliminary proposition. It is the purpose of this section to prove Proposition 13 which states that an R-geodesic consists of pieces, each of which is either a straight line segment or an arc of a circle of radius R. But first we need a few preliminary definitions and propositions.

Let $L(s)$ be a parameterized straight line segment for $a \leqq s \leqq b$, where s is arc length for L. We will say that the line segment *leaves a curve* $X(s)$ *at the point* $X(s_0)$, provided that $L(a) = X(s_0)$ and $L'(a) = X'(s_0)$. Similarly, we say that the parameterized line segment *arrives at the curve* $X(s)$ *at the point* $X(s_0)$, provided that $L(b) = X(s_0)$ and $L'(b) = X'(s_0)$.

We state without proof the following obvious geometric fact.

PROPOSITION 10. *If* $C(s)$ *and* $B(s)$ *are any two distinct similarly oriented parameterized circles of radius R in a plane, then there exists a unique parameterized straight line segment which leaves* $C(s)$ *and arrives at* $B(s)$. *Furthermore if* $C(s)$ *and* $B(s)$ *are oppositely oriented then there*

exists a parameterized straight line segment which leaves $C(s)$ and arrives at $B(s)$ if and only if no point of $B(s)$ is in the interior of $C(s)$. If such a segment exists, then it is unique.

PROPOSITION 11. *Let X be a planar curve with average curvature everywhere less than or equal to R^{-1} defined for $0 \leqq s \leqq d \leqq \pi R/8$. Then the collection $C = C(2, X(0), X'(0), X(d), X'(d), R)$ contains a curve W of type ALA.*

Proof. It is convenient to assume that the plane π which contains the curve X is the collection of all ordered couples of real numbers. It is clearly no loss of generality to assume that $X(0) = (0,0)$ and that $X'(0) = (1,0)$. We consider the four circles Z_0, Z_d, Y_0, Y_d. Clearly Z_0 and Y_0 are the circles of radius R with centers respectively at $(0,R)$ and $(0,-R)$. It is an easy consequence of Proposition 6 that if $X(d)$ is on Y_0 or Z_0, then X is an arc of Y_0 or Z_0 respectively. In this event, X is itself of type *ALA*. Therefore we can assume that $X(d)$ is on neither Z_0 nor Y_0. An application of the techniques of Section 3 shows that the first coordinates of the centers of Z_d and Y_d are strictly positive. There are only three cases to consider: (1) every point on Z_d has a non-negative second coordinate; (2) every point on Y_d has a non-positive second coordinate; (3) some, but not all, points on Z_d have negative second coordinates, and some, but not all, points on Y_d have positive second coordinates. We first consider case (1). Proposition 10 assures the existence of a unique parameterized line segment m which leaves Z_0 and arrives at Z_d. There are now two subcases: (a), the slope of m is less than or equal to the slope of $X'(d)$; and (b), the slope of m is greater than the slope of $X'(d)$. Suppose (a) holds. Let A_1 be the smaller arc of Z_0 whose end points are $(0,0)$ and the point of tangency of m with Z_0. Let A_2 be the smaller arc of Z_d whose end points are $X(d)$ and the point of tangency of m with Z_d. Then one sees that the curve, W, which consists of A_1, followed by m, followed by A_2, satisfies the conclusion of the proposition. Suppose now that subcase (b) holds. Proposition 6 implies that Z_0 and Y_d do not intersect. Therefore Proposition 10 assures the existence of a unique parameterized line segment n which leaves Z_0 and arrives at Y_d. This time we let A_1 be the smaller arc of Z_0 whose end points are $(0,0)$ and the point of tangency of n with Z_0. Similarly, we let A_2 be the smaller arc of Y_d whose end points are $X(d)$ and the point of tangency of n with Y_d. This time one sees that the curve, W, which consists of A_1, followed by n, followed by A_2, satisfies the conclusion of the theorem.

Case (2) is treated similarly. We proceed to case (3). It is clear that the first coordinates of the centers of Z_d and Y_d are unequal. We suppose the first coordinate of the center of Z_d to be less than that of Y_d. The other case can be treated similarly. Propositions 6 and 10 assure the existence of a unique directed line segment m which leaves Y_0 and arrives at Z_d. Let A_1 be the smaller arc of Y_0 whose end points are $(0,0)$ and the point of tangency of m with Y_0. Let A_2 be the smaller arc of Z_d whose end points are $X(d)$ and the point of tangency of m with Z_d. Then one sees that the curve, W, which consists of A_1, followed by m, followed by A_2, satisfies the conclusion of the proposition. This completes the proof.

As an immediate corollary to Propositions 9 and 11 we have:

PROPOSITION 12. *Let X be a planar curve of length less than or equal to $\pi R/8$. Then X is an R-geodesic if and only if X is of type ALA.*

We can now easily establish the main result of this section.

PROPOSITION 13. *Let X be a planar curve defined on a closed finite interval $[0,d]$ parameterized by arc length. Then if X is an R-geodesic, it is a continuously differentiable curve which consists of a finite number of pieces, each of which is either a straight line segment, or an arc of a circle of radius R.*

Proof. Since X' satisfies a Lipschitz condition, X is continuously differentiable. The interval $[0,d]$ can be partitioned into a finite number of subintervals each of length less than or equal to $\pi R/8$. Clearly, X restricted to any of these subintervals is also an R-geodesic. Therefore by the preceding proposition, X, so restricted, is of type ALA. This completes the proof.

6. Principal result. We wish to show that a planar R-geodesic is necessarily a very special curve, i. e. a continuously differentiable curve which consists of at most *three* pieces, each of which is either a straight line segment or an arc of a circle of radius R. Errett Bishop pointed out to the author that in view of Proposition 13 it is sufficient to show that no curve which consists of four such pieces can be an R-geodesic.

Let us designate a continuously differentiable curve by $CCCC$, provided that it consists of precisely four arcs of circles of radius R. Similarly let us designate by $CLCL$ a continuously differentiable curve which consists of precisely four pieces, the first of which is an arc of a circle of radius R, the second a line segment, the third an arc of a circle of radius R, and the last a line segment. Similarly any differentiable curve which consists of

precisely n arcs of circles of radius R and line segments can be represented by an n-tuple of symbols, each of which is either a C or an L.

We already know that there exist planar R-geodesics of type CLC. We will show that there exist planar R-geodesics of type CCC. It is easy to see that any subpath of an R-geodesic is an R-geodesic. We will show that

* *every R-geodesic is necessarily a subpath of a path of type CLC or of type CCC.*

In order to prove that every R-geodesic is necessarily a subpath of a path consisting of three arcs and line segments, it is sufficient to show that no path consisting of four arcs and line segments can be an R-geodesic. There are eight paths of this type, namely $CCCC$, $CCCL$, $CCLC$, $CLCC$, $CLCL$, $LCCC$, $LCCL$ and $LCLC$. If a curve is an R-geodesic, then so is the curve obtained by traversing the path in the opposite direction. Therefore, if we can show that no curve of type $CCLC$ is an R-geodesic, we will also have shown that no curve of type $CLCC$ is an R-geodesic. Thus we wish to show that no curves of type $CCCC$, $CCCL$, $CCLC$, $CLCL$ and $LCCL$, are R-geodesic. Also, if a curve is an R-geodesic, then so is every subpath of the curve. Thus if we can show that no curve of type CCL is an R-geodesic it will follow that neither are curves of type $CCCL$, $CCLC$ and $LCCL$. Likewise, if we can show that no curve of type LCL is an R-geodesic, it will follow that neither is one of type $CLCL$. *Thus in order to show * it is sufficient to show that none of the following three types of curves is an R-geodesic: CCCC, CCL, LCL.* We begin with a proof that no curve of type LCL is an R-geodesic. The proof that no curve of type CCL is an R-geodesic is rather similar, and is therefore omitted. We then show that no curve of type $CCCC$ is an R-geodesic.

LEMMA 1. *If X is a curve of type LCL defined for $0 \leqq s \leqq d$, then it is not an R-geodesic.*

Proof. Let $P_1P_2P_3P_4$ be a curve of type LCL where: (1) P_1P_2 is a line segment, (2) P_2P_3 is an arc of circle of radius R, (3) P_3P_4 is a line segment, (4) P_1P_2 is tangent to P_2P_3 at P_2 and (5) P_3P_4 is tangent to P_2P_3 at P_4.

It is obvious that we may assume P_2P_3 counterclockwise oriented and we do so. We now consider two cases.

Case 1. (See fig. 2.) The length of the arc P_2P_3 is greater than zero but not greater than πR. It is easy to see, by considering subpaths, that we may assume that P_1P_2 and P_3P_4 have the same length. Let S_1 be the counterclockwise oriented circle of radius R, tangent to P_2P_1 at P_1, which is

on the same side of P_1P_2 as is the arc P_2P_3. Similarly, let S_2 be the counterclockwise oriented circle of radius R, tangent to P_3P_4 at P_4 which is on the same side of P_3P_4 as is the arc P_2P_3. There exists a unique line segment Q_1Q_2 which leaves S_1 and arrives at S_2. Since the arc length of

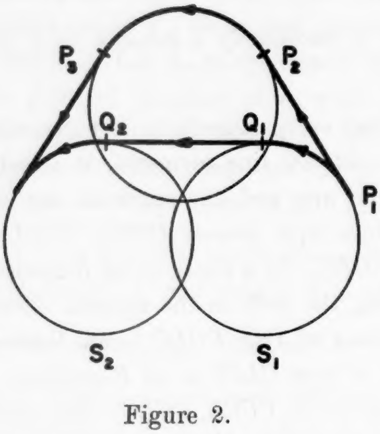

Figure 2.

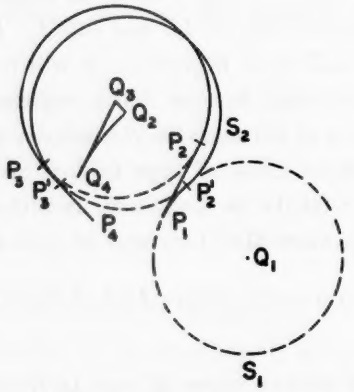

Figure 3.

P_2P_3 does not exceed πR, the curve $P_1Q_1Q_2P_4$ is of type ALA, and hence, by Proposition 9, it is the unique R-geodesic with initial and final positions $X(0)$ *and* $X(d)$ respectively, and initial and final tangent vectors $X'(0)$ and $X'(d)$ respectively.

Case 2. (See fig. 3.) The length of the arc P_2P_3 is strictly between πR and $2\pi R$. It is clear that by considering subpaths we may assume that the segments P_1P_2 and P_3 P_4 do not intersect. Let S_1 *be* the clockwise

oriented circle of radius R tangent to P_1P_2 at P_1 which is on the opposite side of P_1P_2 as is the arc P_2P_3. Let S_2 be the counterclockwise oriented circle of radius R which is tangent to S_1 and to P_3P_4. Let P_2' and P_3' be the points of tangency of S_2 with S_1 and P_3P_4 respectively. We will complete our proof by showing that the path $P_1P_2'P_3'P_4$ has length strictly less than the length of X, where: (a) P_1P_2' is the shorter arc of S_1 with end points P_1 and P_2'; and (b) $P_2'P_3'$ is the longer arc of S_2 with end points P_2' and P_3'; and (c) $P_3'P_4$ is the line segment with end points P_3' and P_4. It is sufficient to show that $P_1P_2'P_3'$ is shorter than $P_1P_2P_3P_3'$, where P_1P_2 and P_2P_3 were as defined above and where P_3P_3' is the line segment with end points P_3 and P_3'. There should be no ambiguity in the remainder of this proof if we use the symbol P_2P_3 to represent the length of the arc P_2P_3 as well as the arc itself, similarly for other arcs and line segments. Thus we will complete our proof by showing

$$P_1P_2' + P_2'P_3' < P_1P_2 + P_2P_3 + P_3P_3'. \tag{1}$$

Let u_2 and u_2' be defined by $Ru_2 = P_2P_3$ and $Ru_2' = P_2'P_3'$. Let u_1' be the angle between the tangent vectors to S_1 at P_1 and P_2'. It is easy to see that (1) is equivalent to

$$P_1P_2' + Ru_2' < P_1P_2 + Ru_2 + P_3P_3'. \tag{2}$$

Furthermore, since $u_2 = u_2' - u_1'$ we see that (2) is equivalent to

$$P_1P_2' + Ru_1' < P_1P_2 + P_3P_3'. \tag{3}$$

For $i = 1$ and 2, let Q_i be the center of S_i. P_2P_3 is an arc of a circle whose center we designate by Q_3. Let Q_4Q_2 be the arc of the circle of radius $2R$ with center Q_1, which is concentric with the arc P_1P_2'. Clearly, the arc Q_4Q_2 has a length equal to the left side of (3), whereas the segments Q_4Q_3 and Q_3Q_2 have lengths equal respectively to P_1P_2 and P_3P_3'. Thus (3) is equivalent to

$$Q_4Q_2 < Q_4Q_3 + Q_3Q_2. \tag{4}$$

Since the arc $Q_4Q_3Q_2$ lies above the convex arc Q_4Q_2, Proposition 7 implies (4). This proves the lemma.

We are indebted to Errett Bishop for the main ideas in the proof of the following lemma. We also wish to thank Horace Moore for pointing out an error in an eariler version.

LEMMA 2. *No path of type CCCC is an R-geodesic.*

Proof. (See fig. 4.) Assume otherwise and let X be a path of type $CCCC$ which is an R-geodesic. Let P_0 and P_4 be the initial and terminal points of X. X consists of four arcs of circles. Let the first arc of X be an arc of the circle S_1 with center Q_1 and let the last arc of X be an arc of the circle S_4 with center Q_4. It is no loss of generality to assume that $R=1$, i.e., the radius of S_1 is 1. Let $2d$ be the distance between Q_1 and Q_4. Clearly $0 \leqq 2d \leqq 6$. It is no loss of generality to assume that S_1 is oriented clockwise. We may also assume that the plane has rectangular coordinates with the origin at Q_1. We may further assume that Q_4 has coordinates

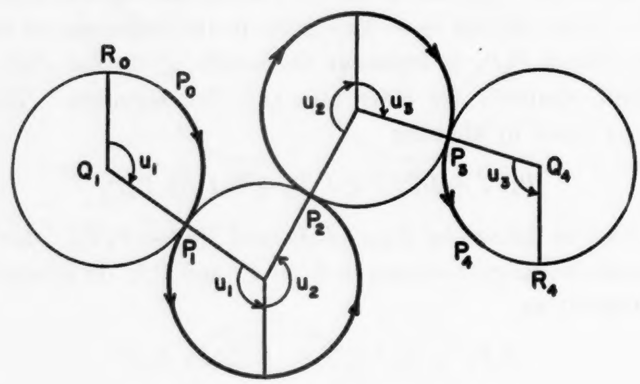

Figure 4.

$(2d, 0)$. Let R_0 be the point $(0,1)$. R_0 is on S_1. Let R_4 be the point $(2d, -1)$. R_4 is on S_4. Now let W be the collection of all paths Y of the type $CCCC$ with initial point P_0 and terminal point P_4 and where the first arc of Y is a clockwise oriented arc of the circle S_1 and where the last arc of Y is a counterclockwise oriented arc of the circle S_4. $Y = P_0P_1P_2P_3P_4$, where $P_{i-1}P_i$ is an oriented arc with length L_i of an oriented circle S_i with center Q_i for $i = 1,2,3,4$ and where S_i is tangent to S_{i+1} at the point P_i for $i = 1,2,3$. Here S_1, S_4, P_0 and P_4 are fixed whereas S_2, S_3, P_1, P_2, P_3, L_1, L_2, L_3 and L_4 vary with Y. Let u_1 be the arc length of the clockwise oriented arc R_0P_1 of the circle S_1. Let u_2 and u_3 be defined by $u_1 + u_2 = L_2$ and $u_2 + u_3 = L_3$. Clearly

$$
\begin{aligned}
L &= L_1 + L_2 + L_3 + L_4 \\
&= (L_1 - u_1 + u_1) + (u_1 + u_2) + (u_2 + u_3) + (u_3 + L_4 - u_3) \\
&= L_1 - u_1 + 2(u_1 + u_2 + u_3) + L_4 - u_3.
\end{aligned}
$$

Furthermore, $L_1 - u_1$ and $L_4 - u_3$ are constants independent of Y in W. Hence, a necessary and sufficient condition for $L(Y)$ to attain its minimum at X is that $u_1 + u_2 + u_3$ attain its minimum at X. Clearly (u_1, u_2, u_3) is an admissible triad if and only if it satisfies the two constraints:

$$(1) \quad \sin u_1 + \sin u_2 + \sin u_3 = \mathrm{d} \quad \text{and} \quad (2) \quad \cos u_1 - \cos u_2 + \cos u_3 = 0.$$

There are now two possibilities for any point, (u_1, u_2, u_3), at which a minimum of $u_1 + u_2 + u_3$, subject to the two constraints, occurs: Either the cross product of the two gradient vectors

$$w_1 = (\cos u_1, \cos u_2, \cos u_3) \quad \text{and} \quad w_2 = (-\sin u_1, \sin u_2, -\sin u_3)$$

is the zero vector or $w_1 \times w_2 \neq 0$. Clearly

$$w_1 \times w_2 = (-\sin(u_2 + u_3), \sin(u_3 - u_1), \sin(u_1 + u_2)).$$

Suppose first that $w_1 \times w_2 = 0$. Then $\sin(u_1 + u_2) = 0$. Then $u_1 + u_2$ is a multiple of π. That is, L_2, the arc length of P_1P_2, is a multiple of π. Since X is assumed to be an R-geodesic of type $CCCC$, it is easy to see that L_2 is neither the zero multiple of π nor can L_2 be as large as 2π. Hence $L_2 = \pi$. Then part of X, namely $P_0P_1P_2P_3$ is a curve of type CCC, where the middle arc is of length π. But we have the following

SUBLEMMA. *No curve of type CCC, where the middle arc is of length greater than zero but not greater than πR, can be an R-geodesic.*

We omit the proof of this sublemma for it is quite similar to the proof of Case 1 of Lemma 1. This sublemma implies that $P_0P_1P_2P_3$ is not an R-geodesic. Hence $X = P_0P_1P_2P_3P_4$ could not be an R-geodesic. Therefore, we need only consider the case $w_1 \times w_2 \neq 0$. Hence, we may assume that the two constraints (1) and (2) determine a curve in the neighborhood of the point (u_1, u_2, u_3) at which the minimum of $u_1 + u_2 + u_3$ is attained, and that $w_1 \times w_2$ is tangent to the curve. Let $m =$ minimum of $u_1 + u_2 + u_3$ subject to the constraints. Then the plane $u_1 + u_2 + u_3 = m$ is tangent to the constraint curve at the point at which the minimum is attained. That is, this plane contains the tangent vector $w_1 \times w_2$ at this point. Equivalently, the vector $(1, 1, 1)$ is orthogonal to $w_1 \times w_2$. Hence

$$-\sin(u_2 + u_3) + \sin(u_3 - u_1) + \sin(u_1 + u_2) = 0.$$

It is not difficult to show that if $\sin(A + B) = \sin A + \sin B$, then either A or B or $A + B$ is a multiple of 2π. Thus, either $u_2 + u_3$ or $u_3 - u_1$ or $u_1 + u_2$ is a multiple of 2π. But $u_1 + u_2 = L_2$, the arc length of P_1P_2.

5

Hence $0 < u_1 + u_2 < 2\pi$. Therefore $u_1 + u_2$ is not a multiple of 2π. Similarly $u_2 + u_3$ is not. Hence $u_3 - u_1$ is a multiple of 2π. Thus at the critical point, $w_1 \times w_2 = (\sin(v_2 + v_1), 0, \sin(v_2 + v_1))$. Since $w_1 \times w_2 \neq 0$, the first component of the tangent $w_1 \times w_2$ is unequal to zero. Therefore, the constraint curve can be parametrized by u_1. at least in a neighborhood of the critical point. In the sequel, all differentiations will be with respect to u_1 in a neighborhood of the critical point. Differentiating the two constraints we get:

$$u_1' \cos u_1 + u_2' \cos u_2 + u_3' \cos u_3 = 0, \qquad -u_1' \sin u_1 + u_2' \sin u_2 - u_3' \sin u_3 =$$

Furthermore, at the critical point, $u_1 + u_2 + u_3$ is a minimum, and hence $u_1' + u_2' + u_3' = 0$. These three equations imply that at a critical point $u_2' \sin(u_1 + u_2) = 0$. Since $\sin(u_1 + u_2) \neq 0$, we conclude that $u_2' = 0$. Hence $u_3' = -u_1' = -1$. We next observe that a necessary condition for $u_1 + u_2 + u_3$ to have a minimum at a regular point of the constraint curve is that the inner product of $r = (1, 1, 1)$ with the curvature vector be nonnegative. Let $Z(s) = (u_1(s), u_2(s), u_3(s))$ be a parametrization of the constraint curve in a neighborhood of the critical point, where s is arc length for the curve. Clearly dz/ds is the tangent vector. We may assume that $w_1 \times w_2(u_1) = b(s) dz/ds$, where $b(s) > 0$ and where $s = s(u_1)$ is a function of u_1. Hence $(w_1 \times w_2)' = b(s)(d^2z/ds^2)s' + (db/ds)s'(dz/ds)$. At the critical point, dz/ds is perpendicular to r. Hence $(r, (w_1 \times w_2)') = b(s)s'(r, d^2z/ds^2)$. Since $b(s) > 0$, we conclude that a necessary condition for $u_1 + u_2 + u_3$ to have a minimum is that $s'(r, (w_1 \times w_2)') \geqq 0$. Clearly

$$\begin{aligned} &(w_1 \times w_2)' \\ &= (-\cos(u_2 + u_3)(u_2' + u_3'), \cos(u_3 - u_1)(u_3' - u_1'), \cos(u_1 + u_2)(u_1' + u_2')). \end{aligned}$$

But at the critical point, $u_1 = u_3$, $u_2' = 0$, $u_1' = 1$, $u_3' = -1$. Hence

$$(w_1 \times w_2)' = (\cos(u_2 + u_1), -2, \cos(u_1 + u_2)).$$

Therefore $(r, (w_1 \times w_2)') = 2\cos(u_1 + u_2) - 1$. Hence, a necessary condition for a minimum is that $2s'(\cos(u_1 + u_2) - 1) \geqq 0$. Hence, a necessary condition for a minimum is either $s' \leqq 0$ or $u_1 + u_2$ a multiple of 2π. Now we observe that the first component of $w_1 \times w_2$ equals the first component of $b(s) dz/ds$. Hence $-\sin(u_2 + u_3) = b(s) du_1/ds$. Since $b(s) > 0$, we have $\sin(u_2 + u_3) \geqq 0$ if and only if $du_1/ds \leqq 0$. Furthermore, since $w_1 \times w_2 \neq 0$ at a regular point, $\sin(u_2 + u_3) \neq 0$ there. Hence

$$\sin(u_2 + u_3) > 0 \iff du_1/ds \leqq 0 \iff ds/du_1 = s' \leqq 0.$$

We again recall that $u_3 = u_1$ at a critical point. Hence, a necessary condition for a minimum of $u_1 + u_2 + u_3$ subject to the two constraints is that $0 \leqq u_1 + u_2 \leqq \pi$. Now $X = P_0P_1P_2P_3P_4$, where L, the arc length of P_1P_2, is $u_1 + u_2$. By the sublemma, $P_0P_1P_2P_3$ is not an R-geodesic. Hence, neither is X. This completes the proof of the lemma.

We have now established our main result:

THEOREM I. *Every planar R-geodesic is necessarily a continuously differentiable curve which is either* (1) *an arc of a circle of radius R, followed by a line segment, followed by an arc of a circle of radius R; or* (2) *a sequence of three arcs of circles of radius R; or* (3) *a subpath of a path of type* (1) *or* (2).

COROLLARY. *There exists an R-geodesic of type CCC.*

Proof. Consider the problem of making a U-turn. That is let X be an R-geodesic whose initial and terminal positions are the same, and whose terminal tangent vector is the negative of its initial tangent vector. Let P be the finite set which consists of the paths of types (1), (2), and (3) referred to in Theorem 1 with the prescribed boundary conditions. In P there are two paths of type CCC. It is at most an elementary calculation to show that no path in P has a length less than either of these. Hence, by Theorem 1, these two paths are R-geodesics. This completes the proof of the corollary.

7. Non-existence of paths of minimal length. Let u, v, U and V be vectors in real n-dimensional Euclidean space, E_n. Let $\| U \| = \| V \| = 1$ and let $R > 0$. Let $C^* = C^*(n, u, U, v, V, R)$ be the collection of all curves X defined on a closed interval $[0, L]$, where $L = L(X)$ varies with X such that: $X(s) \in E_n$ for $0 \leqq s \leqq L$; $\| X'(s) \| = 1$; $X''(s)$ exists everywhere and $\| X''(s) \| \leqq R^{-1}$ for $0 \leqq s \leqq L$; $X(0) = u$, $X'(0) = U$, $X(L) = v$, and $X'(L) = V$.

PROPOSITION 14. *There exist u, U, v, V and R such that the infimum of the length of the curves X in $C^* = C^*(2, u, U, v, V, R)$ is not attained.*

Proof. Let $u = (0,0)$, $U = (0,1)$, $v = (5,0)$, $V = (0,-1)$ and $R = 1$. It is easy to see that there exists a curve Y of type ALA of length $\pi + 3$ such that $Y \in C = C(2, u, U, v, V, R)$. Proposition 9 implies that Y is the unique curve of minimal length in C. It is clear that C^* is a subset of C and that Y is not an element of C^*. To complete the proof of the proposition

it is sufficient to show that for any $\epsilon > 0$, there exists $X \in C^*$ of length less than $\pi + 3 + 4\epsilon$. We will define X uniquely on an interval $[0, L_\epsilon]$ by specifying L_ϵ, $X(0)$, $X'(0)$ and the oriented curvature $k(s)$ for $0 \leqq s \leqq L_\epsilon$. Let $X(0) = u$, $X'(0) = U$, and $k(s) = 1$ for $0 \leqq s \leqq \frac{1}{2}\pi - \epsilon$. Let $k(s)$ be linear for s between $\frac{1}{2}\pi - \epsilon$ and $\frac{1}{2}\pi + \epsilon$, and let $k(\frac{1}{2}\pi + \epsilon) = 0$. Now let d be the first coordinate of the point $X(\frac{1}{2}\pi + \epsilon)$. We can now continue to define k. Let $k(s) = 0$ for $\frac{1}{2}\pi + \epsilon \leqq s \leqq 5 + \frac{1}{2}\pi + \epsilon - 2d$. Let $k(s)$ be linear for s between $5 + \frac{1}{2}\pi + \epsilon - 2d$ and $5 + \frac{1}{2}\pi + 3\epsilon - 2d$. Lastly, let $k(s) = 1$ for $5 + \frac{1}{2}\pi + 3\epsilon - 2d \leqq L_\epsilon$, where $L_\epsilon = 5 + \pi + 2\epsilon - 2d$. Thus X is uniquely defined. Since both the curvature k and its antiderivative can easily be integrated by elementary means, it is at most an elementary calculation to establish that X indeed is in C^*, and that the length L_ϵ of X is indeed less than $\pi + 3 + 4\epsilon$. We omit the details. The intuitive idea is that X was constructed so as to be a twice differentiable curve which approximates Y in an appropriate sense. This completes the proof.

In the course of the proof, we showed that the particular R-geodesic Y in $C(2, u, U, v, V, R)$ had the property that given any $\epsilon > 0$, there exists an X in $C^*(2, u, U, v, V, R)$ whose length differs from the length of Y by a quantity $f(\epsilon)$ which goes to zero with ϵ. It might be conjectured that every R-geodesic has this property. The following is a counterexample to this conjecture. Let Y be a continuously differentiable curve which consists of an arc of length $\frac{1}{2}\pi$ of a counterclockwise oriented circle of radius 1, followed by an arc of length $\frac{1}{2}\pi$ of a clockwise oriented circle of radius 1. We omit the proof that such a Y is indeed a counterexample.

CARNEGIE INSTITUTE OF TECHNOLOGY.

REFERENCES.

[1] L. M. Graves, *The Theory of Functions of Real Variables*, New York and London, 1946.

[2] E. Schmidt, "Über das Extremum der Bogenlänge einer Raumkurve bei vorgeschreibenen Einschränkungen ihrer Krümmung," *Sitzungsberichte Akad. Berlin* (1925), pp. 485-490.

[3] A. Schur, "Über die Schwarzsche Extremaleigenschaft des Kreises unter den Kurven konstanter Krümmung," *Mathematische Annalen*, vol. 83 (1921), pp. 143-148.

RELATIVE CHARACTERISTIC CLASSES.*

By Michel A. Kervaire.

1. Introduction. The main purpose of this paper [1] is to prove a lemma (Lemma (1.2) below) conjectured in [8].

In the proof, we shall make use of a not quite classical form of Whitney duality, involving Stiefel-Whitney characteristic classes which have to be considered as *relative* cohomology classes. Since these slightly generalized characteristic classes may have some interest for themselves, the present paper is divided into two parts as follows.

In Part I, an attempt is made to give a systematic treatment of relative characteristic classes. Beside Stiefel-Whitney classes, relative Chern and Pontryagin characteristic classes will also be considered. It will be seen that most of the properties of the usual characteristic classes may be adapted to hold for the relative classes. In particular, the relative classes satisfy a generalized Whitney duality theorem and Wu's theorem [16] remains true if suitably stated. The fact that Wu's theorem may be extended to the case of a manifold with boundary was communicated to me by R. Thom and was the starting point of the proof of Lemma (1.2). According to R. Thom, this extension of Wu's theorem was first known to H. Cartan, who proved it using (Φ)-cohomology (unpublished). For our purpose, it will be sufficient to reduce (by Lemma (6.1)) the extended Wu's theorem to the ordinary one, thus avoiding (Φ)-cohomology. The proof of the generalized Whitney duality will be based on the interpretation of the relative characteristic classes as symmetric functions. The original author's proof was very cumbersome and will be omitted. The proof given here is due to A. Borel and is reproduced with his permission.

Part II will be concerned with the following situation considered in [8]: Let M_d be a differentiable closed manifold imbedded [2] into some euclidean

* Received January 4,. 1957.

[1] The paper has been completed as the author was under National Science Foundation research contract. The author is also grateful to the Research Commission of Berne University (Switzerland) and to the Swiss Federal School of Technology for grants made available during the preparation of this paper.

[2] All manifolds considered are of class C^∞. Imbedding will mean regular imbedding (and similarly for immersion).

space E_{d+n}, and assume that there exists a continuous field F_n of n-frames (n mutually orthogonal unit vectors) *normal* to M_d in E_{d+n}. To such a pair $(M_d; F_n)$ we attach

(a) a map ω: $M_d \to V_{d+n,n}$ of M_d into the Stiefel manifold of n-frames with base point at the origin in E_{d+n}: the map ω is defined by

$$\omega(x) = \{v_1(x), v_2(x), \cdots, v_n(x)\},$$

where $v_1(x), \cdots, v_n(x)$, are the vectors based at the origin in E_{d+n} and parallel to the vectors of F_n at x; and we also attach

(b) a map f: $S_{d+n} \to S_n$ of the $(d+n)$-dimensional sphere into the n-sphere defined as follows: Take a tubular neighborhood U of M_d in E_{d+n}. Any point $u \in U$ lies in a uniquely determined n-plane N_x normal to M_d at a (uniquely defined) point $x \in M_d$. Using the coordinate system in N_x which is defined by the mutually orthogonal unit vectors of F_n at x, we attach coordinates $y_1, \cdots, y_n$ to the point u. It may be assumed that $\sum_i (y_i)^2 = 1$, if and only if u lies on the boundary of U. The definition of f will involve the mapping

$$r\colon (B_n, \partial B_n) \to (S_n, q^*)$$

of the n-ball B_n onto the n-sphere S_n given (for instance) by the formula

$$(1.1) \quad r(y_1, \cdots, y_n) \\ = (1-2y^2, 2y_1(1-y^2)^{\frac{1}{2}}, 2y_2(1-y^2)^{\frac{1}{2}}, \cdots, 2y_n(1-y^2)^{\frac{1}{2}}),$$

where $y^2 = \sum_k (y_k)^2$ and $q^* \in S_n$ is the point with coordinates $(-1, 0, \cdots, 0)$. Identifying now S_{d+n} with $E_{d+n} + \infty$, the desired mapping f: $S_{d+n} \to S_n$ is given by

$$f(u) = r(y_1, \cdots, y_n) \text{ for } u \in U, \qquad f(x) = q^* \text{ for } x \in S - U.$$

The homotopy class of the map ω (defined under (a) above) is determined[3] by the generalized *curvatura integra* c which represents the homology class of the cycle $\omega(M_d)$ in $H_d(V_{d+n,n}; \mathbf{Z})$. The number c is an integer for d even or $n = 1$ and a remainder mod 2 for d odd $(n > 1)$. Define γ by $\gamma = c - \chi^*(M_d)$, where $\chi^*(M_d)$, the semi-characteristic of M_d, is equal to $\frac{1}{2}\chi(M_d)$ for d even and to $\sum_{i=0}^{s-1} (-1)^i p_i(M_d)$ for d odd, $p_i(M_d)$ being then the rank of $H_i(M_d; \mathbf{Z}_2)$, and $2s = d + 1$.

[3] Recall that the Stiefel manifold of n-frames in $(d+n)$-space is $(d-1)$-connected and its d-dimensional integer homology group is infinite cyclic, or cyclic of order 2, depending on whether d is even or $n = 1$, or d is odd and $n > 1$ respectively.

It has been proved in [8] (Théorème II) that if a manifold M_d' with normal n-frame field F_n' in E_{d+n} leads (by procedure (b)) to a map $f'; S_{d+n} \to S_n$ homotopic to f, then $\gamma = \gamma'$, i.e., $c - \chi^*(M_d) = c' - \chi^*(M_d')$, this formula being valid only mod 2 for d odd.

Actually each homotopy class in $\pi_{d+n}(S_n)$ contains maps $f: S_{d+n} \to S_n$ obtained by the procedure described under (b) and γ is a homomorphism of $\pi_{d+n}(S_n)$ into $\mathbf{Z}$ or $\mathbf{Z}_2$ according as d is even or odd. It has been proved in [8] (Théorème IV) that, for d even, γ is always zero.

Another homotopy invariant associated to f which will play a role in the present paper is Hopf's invariant as generalized by Steenrod [11]: Consider the cell complex $K = S_n \cup e_{d+n+1}$ obtained by attaching the cell e_{d+n+1} to S_n by the given map f, then Steenrod's generalization of Hopf's invariant, which will be denoted by $h(f)$, is the remainder mod 2 defined by

$$Sq^{d+1}(u) = h(f) \cdot v,$$

where u and v are the generators of $H^n(K;\mathbf{Z}_2)$ and $H^{n+d+1}(K;\mathbf{Z}_2)$ respectively and Sq^{q+1} is the Steenrod square which raises the degree by $d+1$.

Each of the invariants $h(f)$ and $\gamma(f)$ has the following properties which can be verified easily (see [8]) for $\gamma(f)$ and were proved by Steenrod ([11], § 18) for $h(f)$:

(1) It is a homomorphism of $\pi_{d+n}(S_n)$ into $\mathbf{Z}_2$ defined for every $d \geqq 1$ and $n \geqq 1$.

(2) It vanishes for d even.

(3) It takes the same value on a homotopy class and on its Freudenthal suspension.

(4) It is zero for every composition map $g \circ f: S_p \to S_r$, where $f: S_p \to S_q$, $g: S_q \to S_r$, provided that $p > q > r$.[4]

(5) If S_d is parallelizable, it takes the value 1 on the standard Hopf's map $s: S_{2d+1} \to S_{d+1}$ with Hopf's invariant 1.

It is not unlikely that properties (1)-(5) should characterize completely $h(f)$. This proves to be true for $d \leqq 7$, even with property (4) omitted.[5] I have no proof of this conjecture for general d. However, using the explicit definitions of $h(f)$ and $\gamma(f)$, we shall prove below the following lemma, which was conjectured in [8]:

LEMMA (1.2). *Using the above notations,* $\gamma(f) = h(f)$.

[4] This fact has not been proved for γ in [8] but is easily seen.

[5] See [8], page 242.

The use which can be made of relative characteristic classes to prove this lemma will become apparent during the proof in Section 8.

I am indebted and wish to express my gratitude to A. Borel and R. Thom for many fruitful discussions during the preparation of the present paper.

PART I. Relative characteristic classes.

2. Definition. We shall treat first Stiefel-Whitney classes. The relative Chern and Pontryagin classes may be obtained similarly and will be discussed briefly in an Appendix (§ 11). The coefficients will be the remainder mod 2, except in § 11 (and § 6).

Let $\mathfrak{B} = (B, p, K, S_{n-1}, \boldsymbol{O}(n))$ be a sphere-bundle over the simplicial complex K with the orthogonal group of n variables $\boldsymbol{O}(n)$ as structural group. Denote by $\mathfrak{B}^q = (B^q, p^q, K, V_{n,n-q}, \boldsymbol{O}(n))$ the bundle associated to $\mathfrak{B}$ with fibre the Stiefel manifold $V_{n,n-q}$ of $(n-q)$-frames (based at a fixed point) in euclidean n-space.

Let L be a subcomplex of K and assume that a cross-section θ^r over L is given in the associated bundle $\mathfrak{B}^r$. This section induces (by the projection $B^r \to B^q$) cross-sections θ^q over L in the associated bundle $\mathfrak{B}^q$ for $q \geqq r$. Roughly speaking, θ^r defines an $(n-r)$-vectorfield $\boldsymbol{F}_{n-r}$ over L, and θ^q is given by the $(n-q)$-vectorfield consisting of the first $(n-q)$ vectors of $\boldsymbol{F}_{n-r}$.

Let W^{q+1} be the $(q+1)$-dimensional Stiefel-Whitney class of the bundle $\mathfrak{B}$. Suppose $q \geqq r$. A representative w^{q+1} of W^{q+1} may be obtained by the usual stepwise extension process over $K^0 \cup L, K^1 \cup L, K^2 \cup L, \cdots, K^q \cup L$ of the cross-section θ^q in $\mathfrak{B}^q$ induced by θ^r over L. The *requirement* that the cross-section over K^q in $\mathfrak{B}^q$ should coincide over L with the given section θ^q leads to a representative w^{q+1} of W^{q+1} which takes the value zero on every $(q+1)$-simplex of L. The cocycle w^{q+1} is thus a representative of a *relative* cohomology class $W_R^{q+1} \in H^{q+1}(K, L; \boldsymbol{Z}_2)$ defined for $q \geqq r$, which will be called the $(q+1)$-dimensional *relative characteristic* Stiefel-Whitney class mod L corresponding to θ^r.

W_R^{q+1} does not depend on the choice of the extension of θ^q over the 0-, 1-, $\cdots$, q-simplexes of K-L (see [12], 33.5). It does, however, depend, in general, on the choice of the given cross-section over L. Let θ^0 and θ^1 be two cross-sections over L in $\mathfrak{B}^r$ and let θ^{0q}, θ^{1q} be their projections in $\mathfrak{B}^q$ $(q \geqq r)$. The stepwise attempt to make the cross-sections θ^{0q} and θ^{1q} coincide meets with an obstruction in dimension q on L. Let $b^q \in H^q(L)$ be the

obstruction cohomology class. It is easily seen, that δb^q is the difference of the $(q+1)$-dimensional Stiefel-Whitney classes corresponding to θ^0 and θ^1, where δ is the coboundary operator of the cohomology sequence:

$$\cdots \leftarrow H^{q+1}(K) \leftarrow H^{q+1}(K,L) \xleftarrow{\delta} H^q(L) \leftarrow H^q(K) \leftarrow \cdots.$$

3. Naturality. Suppose we are given two principal $\boldsymbol{O}(n)$-bundles $\mathfrak{B}$ and $\mathfrak{B}'$ over simplicial complexes K, K' respectively, such that $\mathfrak{B}$ is induced from $\mathfrak{B}'$ by a map $f\colon K \to K'$ which we assume to be simplicial.

Let L and L' be subcomplexes of K and K' respectively, such that $f(L)$ is a subcomplex of L'. Suppose that cross-sections θ^r and θ'^r over L and L' are given in the associated bundles $\mathfrak{B}^r$ and $\mathfrak{B}'^r$ with fibre $V_{n,n-r}$, such that $h\theta^r(x) = \theta'^r f(x)$ for every $x \in L$ (h is the bundle map covering f).

Then, for $q \geqq r$, the relative characteristic Stiefel-Whitney classes $W_R{}^{q+1}$ and $W'_R{}^{q+1}$ of the two bundles are both defined.

LEMMA (3.1). *We have $W_R{}^{q+1} = f^*(W'_R{}^{q+1})$, where f^* is the dual homomorphism $f^*\colon H(K',L';\mathbf{Z}_2) \to H(K,L;\mathbf{Z}_2)$ induced by $f\colon (K,L) \to (K',L')$.* (See [12], 32.7).

Choose an extension over $K'^q \cup L'$ of the cross-section θ'^q in $\mathfrak{B}'^q$ given over L'. Define $W'_R{}^{q+1}$ using this extension. Define θ^q over $K^q \cup L$ as the reverse image by h of the cross-section θ'^q. By assumption, this definition is consistent with the given cross-section over L in $\mathfrak{B}^q$ and the extended θ^q may be used to define $W_R{}^{q+1}$. Let s be a $(q+1)$-simplex of K. If $f(s) = 0$, $|f(s)|$ is a subset of the q-dimensional skeleton of K', thus θ^q may be defined over s and $w^{q+1}(s) = w'^{q+1}(fs) = 0$. If $f(s) \neq 0$, the restriction of f on s is a homeomorphism, and $w^{q+1}(s) = w'^{q+1}(fs)$ follows from $h\theta^q \mid \dot{s} = \theta'^q f \mid \dot{s}$, together with the fact that h induces the identity $h_*\colon \mathbf{Z}_2 \to \mathbf{Z}_2$ (identifying $H_q(V_{n,n-q};\mathbf{Z}_2)$ with $\mathbf{Z}_2$). Thus, for the cohomology classes of w^{q+1} and w'^{q+1}, $W_R{}^{q+1} = f^*(W'_R{}^{q+1})$.

4. The Whitney duality. Let $\mathfrak{B}_i$ be two principal $\boldsymbol{O}(n_i)$-bundles $(i=1,2)$ over the same simplicial complex K, and suppose that cross-sections $\theta_1{}^{r_1}$ and $\theta_2{}^{r_2}$ over subcomplexes L_1, L_2 of K are given in the corresponding associated bundles $\mathfrak{B}_1{}^{r_1}$, $\mathfrak{B}_2{}^{r_2}$.

Let $\mathfrak{B} = \mathfrak{B}_1 \oplus \mathfrak{B}_2$ be the Whitney sum of $\mathfrak{B}_1$ and $\mathfrak{B}_2$. The bundle $\mathfrak{B}$ is an $\boldsymbol{O}(n_1+n_2)$-bundle. The cross-sections $\theta_1{}^{r_1}$, $\theta_2{}^{r_2}$ induce over $L = L_1 \cap L_2$ (where they are both defined) a cross-section θ^r in the associated bundle $\mathfrak{B}^r$

with fibre $V_{n,n-r}$ $(n=n_1+n_2, r=r_1+r_2)$. We shall refer to θ^r as the *sum* of the given cross-sections $\theta_1{}^{r_1}$ and $\theta_2{}^{r_2}$.

Using the sum cross-section θ^r over L, we define the relative Stiefel-Whitney classes $W_R{}^{q+1}$ of $\mathfrak{B} \bmod L$ for $q \geqq r$. Denote by $X_R{}^{k+1}$ and $Y_R{}^{l+1}$ the relative Stiefel-Whitney classes of $\mathfrak{B}_1$ and $\mathfrak{B}_2$ corresponding to the cross-sections $\theta_1{}^{r_1}$, $\theta_2{}^{r_2}$. They are defined for $k \geqq r_1$ and $l \geqq r_2$ respectively.

In the Theorem (4.1) below, expressing Whitney duality for relative classes, the classes $X_R{}^{k+1}$, $Y_R{}^{l+1}$ which are relative classes $\bmod L_1$ and $\bmod L_2$ respectively are regarded as relative classes $\bmod L$. Precisely, we write $X_R{}^{k+1}$. meaning the image of $X_R{}^{k+1}$ by the homomorphism $H^*(K,L_1) \to H^*(K,L)$ induced by the inclusion $(K,L) \to (K,L_1)$, and similarly for $Y_R{}^{l+1}$.

THEOREM (4.1). *The Whitney duality holds for relative characteristic Stiefel-Whitney classes in the following form: For $q \geqq r$, we have*

$$W_R{}^{q+1} = X_R{}^{q+1} + X_R{}^q \cdot Y^1 + \cdots + X_R{}^{r_1+1} \cdot Y^{q-r_1} + \\ + X^{r_1} \cdot Y_R{}^{q-r_1+1} + \cdots + Y_R{}^{q+1}.$$

Notice that some absolute (usual) characteristic classes occur in the right hand side of the above formula. However, in each cup-product $X^a \cdot Y^b$ with $a+b-1=q \geqq r_1+r_2$, either $a-1 \geqq r_1$ or $b-1 \geqq r_2$ (or both), because $a \leqq r_1$ and $b \leqq r_2$ would imply $a+b \leqq r_1+r_2$. Therefore, in each product $X^a \cdot Y^b$, at least one of the classes X^a or Y^b is a relative class and so is every product in the right hand side of (4.1).

The proof of the above theorem will be carried out by showing that the relative characteristic classes may be equivalently defined as symmetric functions (Theorem (5.1)). This alternative definition in turn implies immediately the above duality theorem.

5. Relative characteristic classes as symmetric functions. Proof of Whitney duality. Let us recall Borel's definition of the (usual) characteristic classes [4]. Let $\mathfrak{B}=(E,p,K,\boldsymbol{O}(n))$ be a principal $\boldsymbol{O}(n)$-bundle over K. Let $\boldsymbol{Q}(n)$ denote the subgroup of $\boldsymbol{O}(n)$ consisting of the diagonal matrices $(e_i\delta_{ij})$, with $e_i=+1$ or -1. The space E is also the bundle space of a fibre bundle with fibre $\boldsymbol{Q}(n)=\boldsymbol{O}(1)\times\boldsymbol{O}(1)\times\cdots\times\boldsymbol{O}(1)$ (n factors), which is a covering space since $\boldsymbol{Q}(n)$ is discrete. The base space $\bar{K}=E/\boldsymbol{Q}(n)$ of this bundle is called the space of flags over K. The space $\bar{K}$ is the bundle space of a bundle with fibre $\boldsymbol{F}(n)=\boldsymbol{O}(n)/\boldsymbol{Q}(n)$ and base space K. Let $\rho\colon \bar{K}\to K$ be the projection.

The bundle $\mathfrak{E} = (E, \pi, \bar{K}, \boldsymbol{Q}(n))$ is the Whitney sum of n bundles $\mathfrak{E} = \mathfrak{E}^1 \oplus \mathfrak{E}^2 \oplus \cdots \oplus \mathfrak{E}^n$, where $\mathfrak{E}^i$ is a principal bundle over $\bar{K}$ with structural group $\boldsymbol{O}(1) \cong \boldsymbol{Z}_2$, thus a two-fold covering. Let $x_i \in H^1(\bar{K}; \boldsymbol{Z}_2)$ be the 1-dimensional (only non-zero positive dimensional) Stiefel-Whitney classes of $\mathfrak{E}^i$, $i = 1, 2, \cdots, n$, defined, for instance, as obstruction to the construction of a cross-section in $\mathfrak{E}^i$ over the 1-dimensional skeleton of $\bar{K}$. We have the

THEOREM (A. Borel). *The dual homomorphism $\rho^*: H^*(K) \to H^*(\bar{K})$ is a monomorphism and the ρ^*-image of $H^*(K)$ contains the symmetric functions of the variables $x_1, x_2, \cdots, x_n$.*

It is then legitimate to define the characteristic class W^q of the bundle $\mathfrak{B}$ by the formula $\rho^*(W^q) = S^q(x_1, \cdots, x_n)$, where $S^q(x_1, \cdots, x_n)$ denotes the elementary symmetric function of degree q in the variables $x_1, \cdots, x_n$.

A. Borel has proved in [4] (Théorème 5.1) that the class W^q defined in this way coincides with the characteristic class defined as obstruction.

We come back to relative classes: Suppose a cross-section θ^r is given in the bundle $\mathfrak{B}^r = (E^r, K, V_{n,n-r}, \boldsymbol{O}(n))$ associated to $\mathfrak{B}$ over a subcomplex L of K. Consider the space of flags $\bar{K} = E/\boldsymbol{Q}(n)$ over K. The section θ^r determines a subspace $\bar{L}$ of $\bar{K}$ as follows: Regarding $\mathfrak{B}$ as induced by a map $f: K \to B_{\boldsymbol{O}(n)}$, $f(x)$, $x \in K$, is a non-oriented n-plane in euclidean space of large dimension. A point of E consists of a point $x \in K$, together with an n-frame $v_1, v_2, \cdots, v_n$ in $f(x)$ (see [12], 10.2). Thus a point of $\bar{K}$ may be represented by $\{x; d_1, d_2, \cdots, d_n\}$, where $x \in K$ and $d_1, d_2, \cdots, d_n$ is an ordered set of mutually orthogonal straight lines in $f(x)$. The set $\bar{L}$ consists of those points $\{x; d_1, \cdots, d_n\}$ of $\bar{K}$ such that $x \in L$ and $d_{r+1}, \cdots, d_n$ carry the $(n-r)$ vectors of θ^r.

Let $x_1, \cdots, x_n \in H^1(\bar{K}; \boldsymbol{Z}_2)$ be, as before, the 1-dimensional characteristic classes of the two-fold coverings $\mathfrak{E}^1, \mathfrak{E}^2, \cdots, \mathfrak{E}^n$ over $\bar{K}$, the Whitney sum of which is the bundle $(E, \pi, \bar{K}, \boldsymbol{Q}(n))$. Because $\mathfrak{E}^{r+1}, \cdots, \mathfrak{E}^n$ admit cross-sections over $\bar{L}$ given by θ^r, the $(n-r)$ last x_i, $i = r+1, r+2, \cdots, n$, may be defined as *relative* characteristics classes mod $\bar{L}$ (obstruction to extending over the 1-dimensional skeleton of $\bar{K}$ the cross-section in $\mathfrak{E}^i$, $r+1 \leqq i \leqq n$, already given over $\bar{L}$). Because in any product of $(q+1)$ distinct factors from $x_1, \cdots, x_n$ with $q \geqq r$, at least one of the last $n-r$ must occur, it follows that, for $q \geqq r$, the elementary symmetric function $S^{q+1}(x_1, \cdots, x_n)$ may be (and will be) defined as a relative cohomology class, which we shall denote by $S_R^{q+1}(x_1, \cdots, x_n)$, using the relative $x_{r+1}, \cdots, x_n$.

THEOREM (5.1). (a) *The homomorphism* $\rho_R^*: H^*(K, L) \to H^*(\bar{K}, \bar{L})$ *is a monomorphism and* (b) $\rho_R^*(W_R^{q+1}) = S_R^{q+1}(x_1, \cdots, x_n)$ *for* $q \geqq r$, *where* W_R^{q+1} *is the relative characteristic class* $\operatorname{mod} L$ *corresponding to the cross-section* θ^r.

The proof of this theorem which will be given below is due to A. Borel. In the first formulation of this paper, the point (b) was proved using an inductive argument (on n) and the special case of relative Whitney duality in which one of the bundles involved in the Whitney sum is a two-fold covering. This special case of Whitney duality was proved directly using the definition of characteristic classes as obstruction, which is rather cumbersome. Futhermore, the point (a) of Theorem (5.1) could not be obtained by this method. The Lemma (5.2) below is also due to A. Borel and was unknown to the author.

The Whitney duality formula (4.1) is an immediate consequence of Theorem (5.1) and of the identity

$$S_R^{q+1}(x_1, \cdots, x_m, y_1, \cdots, y_n) = \sum_{a+b=q+1} S^a(x_1, \cdots, x_m) \cdot S^b(y_1, \cdots, y_n),$$

where each product in the right hand sum is a relative class (at least one factor in each product is a relative class if the left hand side is).

Before we proceed to the proof of Theorem (5.1), we give a property of the relative Stiefel-Whitney classes, which will be needed for this proof.

Let L be a subcomplex of $B_{\boldsymbol{O}(n)}$ and θ^r a cross-section over L in the associated bundle with fibre $V_{n,n-r}$. The cross-section θ^r induces cross-sections θ^q over L in every associated bundle with fibre $V_{n,n-r}$ for $q \geqq r$. The bundle space of the associated bundle with fibre $V_{n,n-r}$ is classifying space for $\boldsymbol{O}(q)$ and will consequently be denoted by $B_{\boldsymbol{O}(q)}$. The projection map $B_{\boldsymbol{O}(q)} \to B_{\boldsymbol{O}(n)}$ is the Borel map $\rho(\boldsymbol{O}(q), \boldsymbol{O}(n))$ corresponding to the inclusion $\boldsymbol{O}(q) \to \boldsymbol{O}(n)$. We write $\rho_{q,n}$, meaning $\rho(\boldsymbol{O}(q), \boldsymbol{O}(n))$, for notational convenience.

LEMMA (5.2). *The relative Stiefel-Whitney characteristic class* W_R^{q+1} *is the only non-zero element in* $H^{q+1}(B_{\boldsymbol{O}(n)}, L; \boldsymbol{Z}_2)$ *belonging to the kernel of* $\rho_{q,n}^*: H^*(B_{\boldsymbol{O}(n)}, L) \to H^*(B_{\boldsymbol{O}(q)}, \theta^q L)$.

We first prove that W_R^{q+1} belongs to the kernel of $\rho_{q,n}^*$: The map $\rho_{q,n}: B_{\boldsymbol{O}(q)} \to B_{\boldsymbol{O}(n)}$ induces over $B_{\boldsymbol{O}(q)}$ as base space an $\boldsymbol{O}(q)$-bundle with fibre $V_{n,n-q}$ and the induced cross-section over $\theta^q L$ may be trivially extended all over $B_{\boldsymbol{O}(q)}$. By naturality the Stiefel-Whitney class corresponding to this cross-section is $\rho_{q,n}^*(W_R^{q+1})$. Since the cross-section may be extended, $\rho_{q,n}^*(W_R^{q+1}) = 0$.

In order to prove that W_R^{q+1} is the *only* non-zero element of dimension $q+1$ in the kernel of $\rho_{q,n}{}^*$, let us consider the following diagram, where the rows are the cohomology sequences of the pairs $(B_{\boldsymbol{O}(q)}, \theta^q L)$ and $(B_{\boldsymbol{O}(n)}, L)$ respectively:

$$\begin{array}{ccccccccc}
\cdots \to & H^i(B_{\boldsymbol{O}(q)}) & \xrightarrow{j^*} & H^i(\theta^q L) & \xrightarrow{\delta} & H^{i+1}(B_{\boldsymbol{O}(q)}, \theta^q L) & \longrightarrow & H^{i+1}(B_{\boldsymbol{O}(q)}) & \to \cdots \\
& \uparrow \rho^* & & \downarrow \theta^* & & \uparrow \rho_{q,n}{}^* & & \uparrow \rho^* & \\
\cdots \to & H^i(B_{\boldsymbol{O}(n)}) & \xrightarrow{i^*} & H^i(L) & \xrightarrow{\delta} & H^{i+1}(B_{\boldsymbol{O}(n)}, L) & \xrightarrow{a^*} & H^{i+1}(B_{\boldsymbol{O}(n)}) & \to \cdots
\end{array}$$

(coefficients in $\mathbf{Z}_2$). One has $H^*(B_{\boldsymbol{O}(n)}; \mathbf{Z}_2) = \mathbf{Z}_2[W^1, W^2, \cdots, W^n]$, $H^*(B_{\boldsymbol{O}(q)}; \mathbf{Z}_2) = \mathbf{Z}_2[W^1, W^2, \cdots, W^q]$ and the W^i for $i \leqq q$ correspond to each other by ρ^*. Thus, for $i \leqq q$, the homomorphism ρ^* is an isomorphism.

By the Five Lemma, it follows that for $i \leqq q$, the homomorphism $\rho_{q,n}{}^*$ is also an isomorphism (θ^* is an isomorphism in every dimension).

Let x be an element of the kernel of $\rho_{q,n}{}^*$ in $H^{q+1}(B_{\boldsymbol{O}(n)}, L)$. By commutativity in the above diagram, a^*x is an element of the kernel of ρ^*. Therefore, by [4], Lemma 5.1, $a^*x = c \cdot W^{q+1}$ with $c = 0$ or 1. Thus $y = x + c \cdot W_R^{q+1}$ belongs to the kernels of both $\rho_{q,n}{}^*$ and a^*. By exactness, there exist elements z and t in $H^q(L)$ and $H^q(B_{\boldsymbol{O}(q)})$ respectively, such that $\delta z = y$ and $j^* t = \theta^{*-1} z$. Since ρ^* is an epimorphism in every dimension, there exists a class $w \in H^q(B_{\boldsymbol{O}(n)})$, such that $\rho^* w = t$. It follows that $y = \delta i^* w = 0$, by exactness.

This completes the proof of Lemma (5.2).

Notice that Lemma (5.2) suggests a new (more general) definition of the relative Stiefel-Whitney classes: Let A be a closed subset of $B_{\boldsymbol{O}(n)}$ and suppose a cross-section θ^r over A is given in the bundle with fibre $V_{n,n-r}$ associated to the universal bundle over $B_{\boldsymbol{O}(n)}$. The $(q+1)$-dimension universal relative Stiefel-Whitney class mod A corresponding to θ^r $(r \geqq q)$, may be defined as being the only non-zero element in the kernel of $\rho_{q,n}{}^*$: $H^*(B_{\boldsymbol{O}(n)}, A) \to H^*(B_{\boldsymbol{O}(q)}, \theta^q A)$. The proof of uniqueness runs as in Lemma (5.2), replacing L by A. To complete the definition, one has to show the existence of an element in kernel $\rho_{q,n}{}^*$, the image of which by a^* is the ordinary universal Stiefel-Whitney class and which is thus different from zero. Extension of this definition to the characteristic classes of a principal $\boldsymbol{O}(n)$-bundle over any compact finite dimensional space X is immediate. However, the naturality is less easy to prove in this general case. We omit the details here and shall treat relative Chern classes by this method (see Appendix).

Proof of Theorem (5.1). *Proof of* (a). This part is not concerned with characteristic classes. Let X be the base space (assumed to be a compact finite dimensional topological space) of an $\boldsymbol{O}(n)$-bundle induced by some map $X \to B_{\boldsymbol{O}(n)}$ and let $\bar{X}$ be the space of flags over X (denote by $\rho_X \colon \bar{X} \to X$ the projection). According to [4], Théorème 5.1, the fibre $\boldsymbol{F}(n) = \boldsymbol{O}(n)/\boldsymbol{Q}(n)$ is totally non-homologous to zero in $\bar{X}$ and thus $\rho_X{}^* \colon H^*(X) \to H^*(\bar{X})$ is a monomorphism.

Let A be a closed subspace of X and suppose that a cross-section θ^r over A is given in the associated bundle over X with fibre $V_{n,n-r}$. Let $\bar{A}$ be the subset of $\bar{X}$ the points of which are the sets $\{a; d_1, \cdots, d_n\}$, such that $a \in A$, and $d_{r+1}, \cdots, d_n$ carry the $(n-r)$ vectors given over A by θ^r. Notice that $\bar{A}$ is homeomorphic to the space of the bundle $(\bar{A}, \rho_A, \boldsymbol{F}(r), \boldsymbol{O}(r))$ defined as follows: E being the space of the $\boldsymbol{O}(n)$-bundle over X, consider the fibering $(E/\boldsymbol{Q}(r), \rho_E, E/\boldsymbol{O}(r), \boldsymbol{F}(r), \boldsymbol{O}(r))$, where $\boldsymbol{Q}(r) \subset \boldsymbol{O}(r) \subset \boldsymbol{O}(n)$. The bundle $(\bar{A}, \rho_A, A, \boldsymbol{F}(r), \boldsymbol{O}(r))$ is induced by the cross-section $\theta^r \colon A \to E/\boldsymbol{O}(r)$. Therefore, by Borel's theorem, $\rho_A{}^*$ is also a monomorphism.

Consider the following commutative diagram in which the rows are the cohomology sequences of $(\bar{X}, \bar{A})$ and (X, A) respectively:

$$
(5.3)\qquad
\begin{tikzcd}
\cdots & H^k(\bar{X}) \arrow[l] & H^k(\bar{X}, \bar{A}) \arrow[l, "\bar{j}^*"'] & H^{k-1}(\bar{A}) \arrow[l, "\delta"'] & H^{k-1}(\bar{X}) \arrow[l, "\bar{i}^*"'] & \cdots \arrow[l] \\
\cdots & H^k(X) \arrow[l] \arrow[u, "\rho_X{}^*"] & H^k(X, A) \arrow[l, "j^*"'] \arrow[u, "\rho_R{}^*"'] & H^{k-1}(A) \arrow[l, "\delta"'] \arrow[u, "\rho_A{}^*"'] & H^{k-1}(X) \arrow[l, "i^*"'] \arrow[u, "\rho_X{}^*"'] & \cdots \arrow[l]
\end{tikzcd}
\quad .
$$

We have to prove that $\rho_R{}^*$ is a monomorphism. Let $a \in H^k(X, A)$ be a cohomology class, such that $\rho_R{}^* a = 0$. Since $\bar{j}^* \rho_R{}^* a = \rho_X{}^* j^* a = 0$ and $\rho_X{}^*$ is a monomorphism, it follows that $j^* a = 0$. Thus by exactness, there exists a class $b \in H^{k-1}(A)$, such that $\delta b = a$. Because $\delta \rho_A{}^* b = \rho_R{}^* \delta b = \rho_R{}^* a = 0$, there exists (by exactness) an element $w \in H^{k-1}(\bar{X})$, such that $\bar{i}^* w = \rho_A{}^* b$. The desired conclusion $a = 0$ would be granted if we knew the existence of an element $x \in H^{k-1}(X)$, such that $i^* v = b$. Indeed, from the existence of v follows, by exactness, $a = \delta b = \delta i^* v = 0$.

It remains to prove the

Lemma. *Let $b \in H^*(A)$ and $w \in H^*(\bar{X})$ be cohomology classes, such that $\rho_A{}^* b = \bar{i}^* w$, then there exists a cohomology class $v \in H^*(X)$, such that* $i^* v = b$ (Notations as in diagram (5.3)).

Proof. Let $x_1, \cdots, x_n$ denote, as before, the 1-dimensional characteristic classes of the two-fold coverings $\mathfrak{E}^1, \mathfrak{E}^2, \cdots, \mathfrak{E}^n$ over $\bar{X}$. By definition

of $\bar{A}$, we have $\bar{\imath}^*x_a = y_a$ for $a = 1, 2, \cdots, r$ and $\bar{\imath}^*x_{r+b} = 0$ for $b = 1, \cdots, n-r$, where $y_1, \cdots, y_r$ denote the characteristic classes of the restriction over $\bar{A}$ of the two-fold coverings $\mathfrak{C}^1, \mathfrak{C}^2, \cdots, \mathfrak{C}^r$. The map $\bar{\imath}\colon \bar{A} \to \bar{X}$ restricted on a fibre (we denote this restriction again by $\bar{\imath}$) induces the inclusion $\boldsymbol{F}(r) \to \boldsymbol{F}(n)$ corresponding to the inclusion $(\boldsymbol{O}(r), \boldsymbol{Q}(r)) \to (\boldsymbol{O}(n), \boldsymbol{Q}(n))$. According to the results of A. Borel in [4] (Théorème 11.1), one has

$$H^*(\boldsymbol{F}(n); \mathbf{Z}_2) \cong \mathbf{Z}_2[x_1, \cdots, x_n]/(S^+(x_1, \cdots, x_n))$$

and similarly for $H^*(\boldsymbol{F}(r); \mathbf{Z}_2)$, where $(S^+(x_1, \cdots, x_n))$ denotes the ideal (in $\mathbf{Z}_2[x_1, \cdots, x_n]$) generated by the symmetric functions of positive degrees in the variables $x_1, \cdots, x_n$.

It is easily seen that there exists a basis $h_1, h_2, \cdots, h_t$ of $H^*(\boldsymbol{F}(n))$ over $\mathbf{Z}_2$ with the following properties:

(1) $h_1 = 1$, the h_i are monomials in the $x_1, \cdots, x_n$;

(2) $\bar{\imath}^*h_1, \bar{\imath}^*h_2, \cdots, \bar{\imath}^*h_s$ form a basis of $H^*(\boldsymbol{F}(r); \mathbf{Z}_2)$;

(3) $\bar{\imath}^*h_{s+1} = \bar{\imath}^*h_{s+2} = \cdots = \bar{\imath}^*h_t = 0$.

Such a basis may be obtained by writing down in some order, beginning with 1, all monomials in the variables $x_1, \cdots, x_r$ the degree of which does not exceed $\dim \boldsymbol{F}(r)$, followed by the other monomials in $x_1, \cdots, x_n$. By omitting in this list the monomials which are linearly dependent (modulo the ideal generated by $S^+(x_1, \cdots, x_n)$) of preceding ones, one obtains the desired basis $h_1, h_2, \cdots, h_t$.

By Borel's results, the spectral sequence of the fibering $\rho_X\colon \bar{X} \to X$ is trivial ($E_2 \cong E_\infty$). Furthermore, since we have a coefficient *field*, $\mathbf{Z}_2$, the term E_∞ is additively isomorphic to $H^*(\bar{X})$. Therefore,

$$H^*(\bar{X}) \cong H^*(X) \otimes H^*(\boldsymbol{F}(n))$$

is a module over $\rho_X^*H^*(X)$ with the basis $h_1, h_2, \cdots, h_t$. Similarly, $H^*(\bar{A})$ is a module over $\rho_A^*H^*(A)$ with the basis $\bar{\imath}^*h_1, \bar{\imath}^*h_2, \cdots, \bar{\imath}^*h_s$.

Any element $w \in H^*(\bar{X})$ admits a unique decomposition in the form

$$w = \sum_{1 \leqq a \leqq t} \rho_X^*(v_a) \cdot h_a, \text{ where } v_i \in H^*(X).$$

We have $\bar{\imath}^*w = \sum_{1 \leqq a \leqq s} \bar{\imath}^*\rho_X^*(v_a) \cdot \bar{\imath}^*h_a$, because $\bar{\imath}^*h_a = 0$ for $a = s+1, \cdots, t$.

If $\bar{\imath}^*w = \rho_A^*b$, as we have assumed in the lemma we are proving, then

$$\rho_A^*b = \sum_{1 \leqq a \leqq s} \rho_A^*i^*(v_a) \cdot \bar{\imath}^*h_a.$$

This is (by uniqueness of the representation) only possible if $i^*v_2 = i^*v_3 = \cdots = i^*v_s = 0$. Thus $\rho_A{}^*b = \rho_A{}^*i^*(v_1)$ $(h_1 = 1)$. Therefore, $b = i^*v_1$, since $\rho_A{}^*$ is a monomorphism. This completes the proof of the lemma.

Proof of (b). By naturality, it is sufficient to prove that the relation $\rho_R{}^*(W_R{}^{q+1}) = S_R{}^{q+1}(x_1, \cdots, x_n)$, $q \geqq r$, holds for the bundle

$$(B_{\boldsymbol{Q}(n)}, \rho(\boldsymbol{Q}(n), \boldsymbol{O}(n)), B_{\boldsymbol{O}(n)}, \boldsymbol{F}(n), \boldsymbol{O}(n)),$$

where ρ_R is written for $\rho_R(\boldsymbol{Q}(n), \boldsymbol{O}(n))$. This is easily seen using the fact that the map $f\colon K \to B_{\boldsymbol{O}(n)}$ inducing the given $\boldsymbol{O}(n)$-bundle over K may be approached by a *simplicial injective* map g as soon as $\dim B_{\boldsymbol{O}(n)} \geqq 2 \cdot \dim K + 1$ (for the existence of g, see Theorem 5 in S. Eilenberg, On spherical cycles, Bulletin of the American Mathematical Society, vol. 47 (1941), pp. 432-434).

Let $(B_{\boldsymbol{O}(r)}, \rho_{r,n}, B_{\boldsymbol{O}(n)}, V_{n,n-r}, \boldsymbol{O}(n))$ be the bundle associated to the classifying bundle for $\boldsymbol{O}(n)$, with fibre $V_{n,n-r}$. The bundle space of this bundle is classifying space for $\boldsymbol{O}(r)$ and is consequently denoted by $B_{\boldsymbol{O}(r)}$. The projection $\rho_{r,n}$ is the Borel map $\rho(\boldsymbol{O}(r), \boldsymbol{O}(n))$.

Consider the following commutative diagram

$$(5.4) \qquad \begin{array}{ccc} B_{\boldsymbol{Q}(r)} & \xrightarrow{\mu} & B_{\boldsymbol{Q}(n)} \\ {\scriptstyle \mu^1}\downarrow & & \downarrow \\ B_{\boldsymbol{O}(r)} & \xrightarrow{\rho_{r,n}} & B_{\boldsymbol{O}(n)} \end{array}$$

and let A be a closed subset of $B_{\boldsymbol{O}(n)}$, such that a cross-section θ^r over A is given in the bundle $(B_{\boldsymbol{O}(r)}, \rho_{r,n}, B_{\boldsymbol{O}(n)}, V_{n,n-r}, \boldsymbol{O}(n))$. Let $\bar{A}$ be defined as at the beginning of this section. As noticed previously, $\bar{A}$ is homeomorphic to a subset $\bar{A}^1$ of $B_{\boldsymbol{Q}(r)}$ and $\mu(\bar{A}^1) = \bar{A}$. We have to prove that the element $y = \rho_R{}^*(W_R{}^{q+1}) - S_R{}^{q+1}(x_1, \cdots, x_n) \in H^*(B_{\boldsymbol{Q}(n)}, \bar{A})$ is zero.

Let us consider the following commutative diagram

$$\begin{array}{ccccccccc} \cdots \to & H^q(B_{\boldsymbol{Q}(r)}) & \xrightarrow{\bar{i}^*} & H^q(\bar{A}^1) & \xrightarrow{\delta} & H^{q+1}(B_{\boldsymbol{Q}(r)}, \bar{A}^1) & \longrightarrow & H^{q+1}(B_{\boldsymbol{Q}(r)}) & \to \cdots \\ & \uparrow{\scriptstyle \mu^*} & & \uparrow{\scriptstyle \mu_A{}^*} & & \uparrow{\scriptstyle \mu_R{}^*} & & \uparrow{\scriptstyle \mu^*} & \\ \cdots \to & H^q(B_{\boldsymbol{Q}(n)}) & \xrightarrow{i^*} & H^q(\bar{A}) & \xrightarrow{\delta} & H^{q+1}(B_{\boldsymbol{Q}(n)}, \bar{A}) & \xrightarrow{a^*} & H^{q+1}(B_{\boldsymbol{Q}(n)}) & \to \cdots \end{array}$$

in which the rows are the cohomology sequences of the pairs $(B_{\boldsymbol{Q}(r)}, \bar{A}^1)$ and $(B_{\boldsymbol{Q}(n)}, \bar{A})$ respectively.

In order to prove $y=0$, it is sufficient to prove $\mu_R{}^*(y)=0$ and $a^*(y)=0$. Indeed, if these two equalities hold, there exist elements $z\in H^q(\bar{A})$ and $t\in H^q(B_{Q(r)})$, such that $\bar{i}^*t=\mu_A{}^*z$ and $\delta z=y$. Now,

$$H^*(B_{Q(n)};\mathbf{Z}_2)=\mathbf{Z}_2[x_1,\cdots,x_n],\ H^*(B_{Q(r)};\mathbf{Z}_2)=\mathbf{Z}_2[y_1,\ \cdots,y_r]$$

and $\mu^*x_i=y_i$ for $i=1,2,\cdots,r$, $\mu^*x_{r+j}=0$ for $j=1,\cdots,n-r$. Thus, μ^* is an epimorphism in every dimension. Therefore, there exists an element $w\in H^q(B_{Q(n)})$, such that $\mu^*w=t$. It follows that $y=\delta z=\delta\mu_A{}^{*-1}\bar{i}^*\mu^*w$ $=\delta i^*w=0$.

It remains to prove that $\mu_R{}^*y=0$ and $a^*y=0$.

Proof of $a^*y=0$. This is obvious with regard to the corresponding result of Borel on absolute characteristic classes:

$$a^*y=a^*\rho_R{}^*W_R{}^{q+1}-a^*S_R{}^{q+1}(x_1,\cdots,x_n)=\rho^*W^{q+1}-S^{q+1}(x_1,\cdots,x_n)=0$$

Proof of $\mu_R{}^*y=0$. We prove separately that $\mu_R{}^*\rho_R{}^*W_R{}^{q+1}=0$ and $\mu_R{}^*S_R{}^{q+1}(x_1,\cdots,x_n)=0$.

The first assertion follows from $\mu_R{}^*\rho_R{}^*W_R{}^{q+1}=\mu_1{}^*\rho_{r,n}{}^*W_R{}^{q+1}$ (see diagram (5.4)), and $\rho_{r,n}{}^*W_R{}^{q+1}=0$ proved in Lemma (5.2).

The second assertion, $\mu_R{}^*S_R{}^{q+1}(x_1,\cdots,x_n)=0$ follows from $\mu^*x_{r+j}=0$ for $j=1,2,\cdots,n-r$ (and thus $\mu_R{}^*x_{r+j}=0$, since $a^*\colon H^1(B_{Q(r)},\bar{A}^1)\to H^1(B_{Q(r)})$ is a monomorphism, A being assumed to be non-void) and the fact that for $q\geqq r$, each product of $(q+1)$ distinct factors from $x_1,\cdots,x_n$ must contain at least one x_{r+j} with $1\leqq j\leqq n-r$. Thus each product in $S_R{}^{q+1}(x_1,\cdots,x_n)$ is mapped into zero by $\mu_R{}^*$.

This completes the proof of Theorem (5.1).

6. A lemma on Lefschetz-Poincaré duality. Let $\boldsymbol{G}_1$, $\boldsymbol{G}_2$, $\boldsymbol{G}$ be coefficient groups (abelian) with a pairing $\boldsymbol{G}_1\times\boldsymbol{G}_2\to\boldsymbol{G}$ of the two first groups to the third.

Let (X,A) be an admissible pair for cohomology theory. Assume X to be connected and A to be a neighborhood retract in X (hence the excision $e\colon(X,A)\to(Y,X')$ induces an isomorphism $H^*(Y,X')\to H^*(X,A)$).

Let Y be the space obtained by matching together two copies of X along the copies of A (i.e., $Y=X+X'$, with A and A' pointwise identified).

Let n be some positive integer. Denote by ${}_1H$, ${}_2H$, H cohomology groups with coefficients in $\boldsymbol{G}_1$, $\boldsymbol{G}_2$, $\boldsymbol{G}$ respectively.

LEMMA (6.1). *Assume $H^n(X)=0$, then the pairings of ${}_1H^q(X)$ with ${}_2H^{n-q}(X,A)$ to $H^n(X,A)$ and of ${}_1H^q(X,A)$ with ${}_2H^{n-q}(X)$ to* $\boldsymbol{H}^n(X,A)$

given by the cup-product are completely orthogonal if and only if the pairing of ${}_1H^q(Y)$ *with* ${}_2H^{n-q}(Y)$ *to* $H^n(Y)$ *is completely orthogonal* (q, n fixed, $0 \leqq q \leqq n$).

Recall that a pairing is said to be completely orthogonal if either of the first two groups involved is the group of all homomorphisms of the other into the third.

Proof. Consider for $a = 1, 2$ the cohomology sequence of the pair (Y, X)

$$\cdots \to {}_aH^{q-1}(X) \xrightarrow{\delta} {}_aH^q(Y, X) \xrightarrow{h^*} {}_aH^q(Y) \xrightarrow{i^*} {}_aH^q(X) \to \cdots.$$

Since (by excision) the inclusion map $e\colon (X', A') \to (Y, X)$ induces isomorphisms $e^*\colon {}_aH^q(Y, X) \to {}_aH^q(X', A')$, we may substitute ${}_aH^q(X', A')$ for ${}_aH^q(Y, X)$ in the above sequence. Specifically, we consider the exact sequence

$$\cdots \to {}_aH^{q-1}(X) \xrightarrow{\delta'} {}_aH^q(X', A') \xrightarrow{j'^*} {}_aH^q(Y) \xrightarrow{i^*} {}_aH^q(X) \to \cdots,$$

where $\delta' = e^*\delta$ and $j'^* = h^*e^{*-1}$.

Let $k\colon Y \to X$ be the map defined by $k(x) = x$, $k(x') = x$, where x' corresponds to x in the copy X' of X. One has $ki = \text{id.}$, and therefore $i^*k^* = \text{id.}$. Thus i^* is an epimorphism, k^* a monomorphism and δ' is trivial. The sequence

$$0 \to {}_aH^q(X', A') \xrightarrow{j'^*} {}_aH^q(Y) \xrightarrow{i^*} {}_aH^q(X) \to 0$$

is exact.

Moreover, ${}_aH^q(Y)$ is the direct sum

$$(6.2) \qquad {}_aH^q(Y) = j'({}_aH^q(X', A')) + k({}_aH^q(X))$$

(we drop the stars by j'^* and k^* for notational convenience), or alternatively, interchanging X and X':

$$(6.3) \qquad {}_aH^{n-q}(Y) = k'({}_aH^{n-q}(X')) + j({}_aH^{n-q}(X, A)).$$

Let us denote by $l'\colon {}_aH^q(Y) \to {}_aH^q(X', A')$, respectively $l\colon {}_aH^q(Y) \to {}_cH^q(X, A)$ the homomorphisms, such that $l'j' = id.$, and $lj = id.$.

We have, for every $u_1 \in {}_1H^q(Y)$ and $x \in {}_2H^{n-q}(X, A)$,

$$(6.4) \qquad j(iu_1 \cdot x) = u_1 \cdot jx,$$

which may be proved using 3.4 of [11]. Setting $u_1 = ka$ and $x = ly$, one obtains

(6.5) $$l(ka \cdot y) = a \cdot ly.$$

Similarly,

(6.5′) $$l'(b \cdot k'x') = l'b \cdot x'.$$

1. *Assume the pairings of* ${}_1H^q(X)$ *with* ${}_2H^{n-q}(X,A)$ *to* $H^n(X,A)$ *and of* ${}_1H^q(X,A)$ *with* ${}_2H^{n-q}(X)$ *to* $H^n(X,A)$ *to be completely orthogonal.*

(1a) Let $h: {}_2H^{n-q}(Y) \to H^n(Y)$ be a homomorphism. Define the homomorphisms $g: {}_2H^{n-q}(X,A) \to H^n(X,A)$ and $g': {}_2H^{n-q}(X') \to H^n(X',A')$ by $g(x) = lhjx$ and $g'(x') = l'hk'x'$ respectively. By assumption, there exist elements $a \in {}_1H^q(X)$ and $a' \in {}_1H^q(X',A')$, such that $g(x) = a \cdot x$ and $g'(x') = a' \cdot x'$ for every $x \in {}_2H^{n-q}(X,A)$ and $x' \in {}_2H^{n-q}(X')$. Set $b = ka + j'a'$ (thus $ib = a$, $l'b = a'$). We have $h(y) = b \cdot y$ for every $y \in {}_2H^{n-q}(Y)$. Indeed, each y admits a decomposition $y = k'x' + jx$ and $lhjx = g(x) = a \cdot x$, $l'hk'x' = g'(x') = a' \cdot x'$.

Therefore, $h(y) = hjx + hk'x' = j(a \cdot x) + j'(a' \cdot x')$. Notice that

$$l'(b \cdot k'x') = l'b \cdot x' = a' \cdot x'$$

(by (6.5′)). We obtain

$$h(y) = j(ib \cdot x) + j'l'(b \cdot k'x') = b \cdot jx + b \cdot k'x' = b \cdot (jx + k'x') = b \cdot y$$

($j'l' = \text{id.}$ in dimension n follows from $H^n(X) = 0$, because then $l': H^n(Y) \to H^n(X',A')$ is an isomorphism).

(1b) Suppose that $b \cdot y = 0$ for some $b \in {}_1H^q(Y)$ and every $y \in {}_2H^{n-q}(Y)$. We have to prove $b = 0$. By (6.2), b admits a decomposition $b = j'a' + ka$. We have $j(a \cdot x) = j(ib \cdot x) = b \cdot jx = 0$, furthermore, $a' \cdot x' = l'(b \cdot k'x') = l'(0) = 0$ for every $x \in {}_2H^{n-q}(X,A)$ and $x' \in {}_2H^{n-q}(X')$. By assumption, it follows that $a = a' = 0$. Thus, $b = 0$.

(1a) and (1b) prove that the pairing of ${}_1H^q(Y)$ with ${}_2H^{n-q}(Y)$ to $H^n(Y)$ is completely orthogonal.

2. *Assume now the pairing of* ${}_1H^q(Y)$ *with* ${}_2H^{n-q}(Y)$ *to* $H^n(Y)$ *to be completely orthogonal.*

(2a) Take a homomorphism $h: {}_2H^{n-q}(X,A) \to H^n(X,A)$. By (6.3), each $y \in {}_2H^{n-q}(Y)$ may be written uniquely in the form $y = k'x' + jx$. The map $g: {}_2H^{n-q}(Y) \to H^n(Y)$ defined by $g(y) = jh(x)$ is a homomorphism and thus, by assumption, there exists an element $u_1 \in {}_1H^q(Y)$, such that $u_1 \cdot y = jh(x)$ for every $y = k'x' + jx \in {}_2H^{n-q}(Y)$. One has (with $u = iu_1$),

$j(u \cdot x) = j(iu_1 \cdot x) = u_1 \cdot jx = jh(x)$. Since j is a monomorphism, $u \cdot x = h(x)$ for every $x \in {}_2H^{n-q}(X, A)$.

(2b) Let $a \cdot x = 0$ for some $a \in {}_1H^q(X)$ and every $x \in {}_2H^{n-q}(X, A)$. Then $a = 0$. Indeed, consider $ka = b \in {}_1H^q(Y)$. We have $b \cdot y = 0$ for every $y \in {}_2H^{n-q}(Y)$, because $l(b \cdot y) = l(ka \cdot y) = a \cdot ly = 0$. In dimension n, l is a monomorphism, because $H^n(X) = 0$. The pairing of ${}_1H^q(Y)$ with ${}_2H^{n-q}(Y)$ to $H^n(Y)$ being assumed to be completely orthogonal, it follows that $b = 0$. Thus, $a = ika = ib = 0$.

The proofs of (2a′): every homomorphism ${}_2H^{n-q}(X) \to H^n(X, A)$ may be "realized" by cup-product with an element of ${}_1H^q(X, A)$ and of (2b′): if $a \cdot x = 0$ for some $a \in {}_1H^q(X, A)$ and every $x \in {}_2H^{n-q}(X)$, then a $= 0$, are mechanical and similar to (2a) and (2b) and will be omitted.

Remark. The assumption $H^n(Y) = 0$ is actually needed in the proofs of both (1) and (2) as is shown by the following examples.

(1) Let $X = P_5$ be the 5-dimensional real projective space and A a point of P_5. Let $\boldsymbol{G}_1 = \boldsymbol{G}_2 = \boldsymbol{G} = \boldsymbol{Z}_2$ and take $n = 5$, $q = 2$. Then ${}_1H^2(X) \cong {}_2H^3(X, A) \cong H^5(X, A) \cong \boldsymbol{Z}_2$ and the pairing of ${}_1H^2(X)$ with ${}_2H^3(X, A)$ to $H^5(X, A)$ by cup-product is completely orthogonal. Similarly, the pairing of ${}_1H^2(X, A)$ with ${}_2H^3(X, A)$ is also completely orthogonal.

Now $Y = P_5 \vee P_5$ and ${}_1H^2(Y) \cong {}_2H^3(Y) \cong H^5(Y) \cong \boldsymbol{Z}_2 + \boldsymbol{Z}_2$. However, $\mathrm{Hom}(\boldsymbol{Z}_2 + \boldsymbol{Z}_2, \boldsymbol{Z}_2 + \boldsymbol{Z}_2) \cong \boldsymbol{Z}_2 + \boldsymbol{Z}_2 + \boldsymbol{Z}_2 + \boldsymbol{Z}_2$. One has $H^5(X) \cong \boldsymbol{Z}_2$.

(2) Let $X = A = S_1$. Then $Y = X = S_1$. Take $n = 1$, $q = 0$ and $\boldsymbol{G}_1 = \boldsymbol{G}_2 = \boldsymbol{G} = \boldsymbol{Z}$. The pairing of ${}_1H^0(Y) = \boldsymbol{Z}$ with ${}_2H^1(Y) = \boldsymbol{Z}$ to $H^1(Y) = \boldsymbol{Z}$ is completely orthogonal, but since ${}_2H^1(X, A) = 0$, $H^1(X, A) = 0$, the pairing of ${}_1H^0(X) = \boldsymbol{Z}$ with ${}_2H^1(X, A)$ to $H^1(X, A)$ is not. One has $H^1(X) = \boldsymbol{Z}$.

7. The relative Wu classes. We come back to coefficients in the field $\boldsymbol{Z}_2$.

Let K be a complex of dimension n and L a non-void subcomplex of K. Suppose that relative Lefschetz-Poincaré duality holds in $K \bmod L$. In other words, for every $q = 0, 1, \cdots, n$, the pairing of $H^q(K; \boldsymbol{Z}_2)$ with $H^{n-q}(K, L; \boldsymbol{Z}_2)$ to $H^n(K, L; \boldsymbol{Z}_2)$ is completely orthogonal. Then we may define Wu classes $U^q \in H^q(K; \boldsymbol{Z}_2)$ by the requirement that for every relative class $X_R^{n-q} \in H^{n-q}(K, L)$

$$Sq^q(X_R^{n-q}) = U^q \cdot X_R^{n-q} \tag{7.1}$$

should hold.

Because L is non-empty, $H^0(K, L) = 0$, and thus, by duality, $H^n(K) = 0$. Consequently, according to Lemma (6.1), absolute Poincaré duality holds in $M = K + K'$ (L and L' identified). Let S^q be the q-dimensional Wu class of M (in the ordinary sense, see [16]), i.e.,

$$Sq^q(X^{n-q}) = S^q \cdot X^{n-q} \tag{7.2}$$

for every class $X^{n-q} \in H^{n-q}(M)$.

LEMMA (7.3). *Let $i\colon K \to M$ be the inclusion map and i^* the dual homomorphism induced by i, then $U^q = i^* S^q$.*

Proof. Let $j^*\colon H^*(K, L) \to H^*(M)$ be as in (6.3). We have, using (6.4), $j^*(i^*S^q \cdot X_R^{n-q}) = S^q \cdot j^* X_R^{n-q} = Sq^q(j^* X_R^{n-q}) = j^*(Sq^q(X_R^{n-q}))$. Since j^* is a monomorphism and the class U^q is determined uniquely by (7.1), it follows that $U^q = i^* S^q$.

Suppose now a sphere-bundle is given over M such that its Stiefel-Whitney characteristic classes $W_M{}^q$ be connected to the Wu classes by the relation

$$W_M{}^q = \sum_{0 \leq p \leq q} Sq^{q-p}(S^p), \tag{7.4}$$

which we may write more conveniently as $W_M = Sq(S)$, denoting by W_M and S the "total" classes, i.e.,

$$W_M = 1 + W_M{}^1 + \cdots + W_M{}^n, \qquad S = 1 + S^1 + \cdots + S^n$$

and where the operator Sq stands for $Sq = Sq^0 + Sq^1 + \cdots + Sq^k + \cdots$.

According to a theorem of Wu in [16], the situation described by formula (7.4) arises in particular if M is a closed differentiable manifold and the sphere-bundle considered is its tangent bundle. From the relative Lefschetz-Poincaré duality for manifolds with regular boundary ([11], 7), it follows that (7.4) *holds in particular with $M = K + K'$ if K is a manifold with regular boundary L.*

Since the characteristic classes of the restricted bundle over K are the i^*-images of the characteristic classes of the bundle over M (by naturality), we have $W = Sq(U)$, where W denotes the total Stiefel-Whitney class of the bundle over K.

Suppose now that

(7.5) *a cross-section over the subcomplex L is given in the associated principal bundle.*

Then, for every q, W^q admits a representative cocycle which vanishes on L and thus defines a relative class $W_R{}^q$.

Since $H^0(K,L)=0$, the algebra $H^*(K,L)$ has no unit element. It is convenient to consider, rather than $H^*(K,L)$, the direct sum $H_1^*=\mathbf{Z}_2+H^*(K,L)$. In other words, we introduce formally a unit into the algebra $H^*(K,L)$. The requirement $1\cdot x=x\cdot 1=x$ for every $x\in H_1^*$ gives to H_1^* a ring structure. We shall furthermore allow the Steenrod squares to operate (as homomorphisms again) in H_1^* by setting $Sq^0(1)=1$, $Sq^i(1)=0$ for $i>0$.

We may then use the classes $W_R{}^q$ to define a total relative class by $W_R=1+W_R{}^1+\cdots+W_R{}^n\in H_1^*$. Since the endomorphism $Sq\colon H_1^*\to H_1^*$ defined for each $X\in H_1$ by $Sq(X)=Sq^0(X)+Sq^1(X)+\cdots$ is a monomorphism and therefore maps H_1^*, as a finite dimensional vector space, *onto* itself, we may define *relative Wu classes* $U_R{}^q$ by

$$W_R=Sq(U_R), \tag{7.6}$$

U_R being the total class $U_R=1+U_R{}^1+U_R{}^2+\cdots\in H_1^*$.

Notice furthermore, that according to the above conventions, we have $Sq(1)=1$, and therefore the product formula of H. Cartan

$$Sq(X\cdot Y)=Sq(X)\cdot Sq(Y) \tag{7.7}$$

holds also in H_1^*.

We state now some properties of the relative Wu class U_R.

Lemma (7.8). *Let $h\colon(K,0)\to(K,L)$ be the inclusion map and h^* the induced homomorphism, then $U^q=h^*U_R{}^q$ for every $q>0$.*

The proof is immediate: Apply h^* to both sides of the equation (7.6) $W_R=Sq(U_R)$, with the convention $h^*(1)=1$. We obtain $W=Sq(h^*U_R)$. Since the class U is uniquely determined by $W=Sq(U)$, Sq being an automorphism of $H^*(K)$, it follows that $U=h^*U_R$.

Lemma (7.9). *For every relative class $X_R{}^{n-q}\in H^{n-q}(K,L)$, we have $Sq^q(X_R{}^{n-q})=U_R{}^q\cdot X_R{}^{n-q}$.*

This is again obvious, according to the preceding lemma.

The following is a discrepancy between properties of absolute and relative Wu classes: Absolute classes the dimension of which exceed $\frac{1}{2}n$ vanish (if $q>\frac{1}{2}n$, then $q>n-q$; thus, in formula (7.1), the square is zero for every $X_R{}^{n-q}$ and, by duality U^q, must be zero, too). This need *not* be the case for relative Wu classes.

We shall be mainly interested in the sequel in the case *n even.* We state some lemmas in this case:

LEMMA (7.10). *Let n be even, $n=2s$. Under assumptions* (7.4) *and* (7.5), *we have $W_R{}^n = U_R{}^s \cdot U_R{}^s + U_R{}^{2s}$.*

Proof.

$$\begin{aligned} W_R{}^n &= Sq^s(U_R{}^s) + Sq^{s-1}(U_R{}^{s+1}) + \cdots + Sp^{s-i}(U_R{}^{s+i}) + \cdots + U_R{}^{2s} \\ &= U_R{}^s \cdot U_R{}^s + U_R{}^{s-1} \cdot U_R{}^{s+1} + \cdots + U_R{}^{s-i} \cdot U_R{}^{s+i} + \cdots + U_R{}^{2s} \\ &= U_R{}^s \cdot U_R{}^s + Sq^{s+1}(U_R{}^{s-1}) + \cdots + Sq^{s+i}(U_R{}^{s-i}) + \cdots + U_R{}^{2s} \\ &= U_R{}^s \cdot U_R{}^s + U_R{}^{2s}. \end{aligned}$$

LEMMA (7.11). *Suppose K has even dimension $n=2s$. Let r be the rank of the bilinear form $f(X,Y)$ over $\mathbf{Z}_2$ defined for $X, Y \in H^s(K,L;\mathbf{Z}_2)$ by $X \cdot Y = f(X,Y)A$, where A denotes the generator of $H^n(K,L)$. Then, under assumptions* (7.4) *and* (7.5) $r \cdot A = U_R{}^s \cdot U_R{}^s$.

Proof. Introduce in $H^s(K,L)$ a basis $Z_1, \cdots, Z_r, Z_{r+1}, \cdots, Z_p$, such that $Z_i \cdot Z_j = \delta_{ij}A$ for $1 \leqq i, j \leqq r$ and $Z_i \cdot Z_j = 0$ if $r < i$ or $r < j$.

With respect to such a basis, $U_R{}^s$ must have the form

$$U_R{}^s = Z_1 + Z_2 + \cdots + Z_r + c_{r+1}Z_{r+1} + \cdots + c_pZ_p.$$

Indeed, let $U_R{}^s = \sum_{1 \leqq i \leqq p} c_iZ_i$, with $c_i \in \mathbf{Z}_2$, and let $X = \sum_{1 \leqq i \leqq p} x_iZ_i$ be any class in $H^s(K,L)$. One has

$$Sq^s(X) = X \cdot X = x_1{}^2 + x_2{}^2 + \cdots + x_r{}^2 = x_1 + x_2 + \cdots + x_r \quad (\text{mod } 2),$$

$$U_R{}^s \cdot X = c_1x_1 + c_2x_2 + \cdots + c_rx_r.$$

Thus, for any choice of $x_1, \cdots, x_r$, Lemma (7.9) implies

$$x_1 + x_2 + \cdots + x_r = c_1x_1 + c_2x_2 + \cdots + c_rx_r.$$

This is only possible if $c_1 = c_2 = \cdots = c_r = 1$.

Now from $U_R{}^s = Z_1 + Z_2 + \cdots + Z_r + c_{r+1}Z_{r+1} + \cdots + c_pZ_p$, it follows that $U_R{}^s \cdot U_R{}^s = Z_1 \cdot Z_1 + Z_2 \cdot Z_2 + \cdots + Z_r \cdot Z_r = rA$, and the proof of (7.11) is complete.

We consider now the dual classes $\bar{W}_R$ and $\bar{U}_R$ which are defined by

$$(7.12) \qquad W_R \cdot \bar{W}_R = 1, \qquad U_R \cdot \bar{U}_R = 1$$

respectively. $\bar{W}_R$ and $\bar{U}_R$ are uniquely defined (as relative classes) because the cup-product with a total class U_R or W_R defines an *isomorphism* $H_1{}^* \to H_1{}^*$

of H_1^* onto itself. Of course, $\bar{W} = h^*\bar{W}_R$ and similarly with U substituted for W.

We have

$$\bar{W}_R = Sq(\bar{U}_R). \tag{7.13}$$

Proof. Denoting $Sq(\bar{U}_R)$ by $\tilde{W}_R = Sq(\bar{U}_R)$, we have

$$W_R \cdot \tilde{W}_R = Sq(U_R) \cdot Sq(\bar{U}_R) = Sq(U_R \cdot \bar{U}_R) = Sq(1) = 1.$$

Since $W_R \cdot \bar{W}_R = 1$ determines $\bar{W}_R$ uniquely, $\tilde{W}_R = \bar{W}_R$.

LEMMA (7.14). *We have (with $n = 2s$ and under assumptions* (7.4), (7.5)) $rA^n = W^{n-1} \cdot \bar{W}_R^1 + W^{n-2} \cdot \bar{W}_R^2 + \cdots + W^1 \cdot \bar{W}_R^{n-1}$, *where* $W^q = h^*W_R^q$ *as in* (7.8); r *and* A^n *were defined in* (7.11).

Proof. We prove first $\bar{W}_R^n = U_R^n$, as follows:

$$\begin{aligned}\bar{W}_R^n &= Sq^0(\bar{U}_R^n) + Sq^1(\bar{U}_R^{n-1}) + \cdots + Sq^s(\bar{U}_R^s)\\ &= \bar{U}_R^n + U_R^1 \cdot \bar{U}_R^{n-1} + \cdots + U_R^s \cdot \bar{U}_R^s\\ &= U_R^{s+1} \cdot \bar{U}_R^{s-1} + \cdots + U_R^{n-1} \cdot \bar{U}_R^1 + U_R^n \quad (\text{by } (7.12))\\ &= Sq^{s+1}(\bar{U}_R^{s-1}) + \cdots + Sq^{n-1}(\bar{U}_R) + U_R^n = U_R^n.\end{aligned}$$

From $W_R^n + W_R^{n-1} \cdot \bar{W}_R^1 + \cdots + W_R^1 \cdot \bar{W}_R^{n-1} + \bar{W}_R^n = 0$, we obtain

$$W_R^n = W_R^{n-1} \cdot \bar{W}_R^1 + \cdots + W_R^1 \cdot \bar{W}_R^{n-1} + U_R^n.$$

Comparing this formula with (7.10) and (7.11), we see that

$$rA^n = W_R^{n-1} \cdot \bar{W}_R^1 + \cdots + W_R^1 \cdot \bar{W}_R^{n-1},$$

hence $rA^n = W^{n-1} \cdot \bar{W}_R^1 + \cdots + W^1 \cdot \bar{W}_R^{n-1}$. This completes the proof of Lemma (7.14).

PART II. Proof and Consequences of Lemma (1.2).

8. The proof. All homology and cohomology groups occurring in this section will be based on remainders mod 2 as coefficients (we shall therefore omit to mention the coefficient field explicitly).

Consider the situation described in Section 1: a C^∞-d-manifold M_d regularly imbedded in euclidean $(d+n)$-space E_{d+n} with a continuous field of normal n-frames $\boldsymbol{F}_n$. In order to prove the Lemma (1.2), i.e., $\gamma = h$, it is sufficient to consider the special case $n = d+1$. Indeed, if $\gamma(f) = h(f)$ has been proved for every $f \in \pi_{2d+1}(S_{d+1})$, the general assertion follows from the

fact that γ and h are both "stable" by suspension (Section 1, property (3)): If $f \in \pi_{d+n}(S_n)$, with $n \leqq d+1$, then $\gamma(f) = \gamma(E^{d+1-n}f) = h(E^{d+1-n}f) = h(f)$. If $f \in \pi_{d+n}(S_n)$, with $n \geqq d+1$, then there exists by Freudenthal's theorems a map $g \in \pi_{2d+1}(S_{d+1})$, such that $E^{n-d-1}g = f$, and for the same reason $\gamma(f) = \gamma(g) = h(g) = h(f)$.

In the sequel $n = d+1$. Let $f\colon S_{2d+1} \to S_{d+1}$ be the sphere map corresponding to the given manifold M_d in E_{2d+1}, together with the n-field $\boldsymbol{F}_n$ $(n = d+1)$ of mutually orthogonal unit vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_n$ normal to M_d in E_{d+n} (see Section 1, (b)). It is easily seen from the definition of f (see (1.1)) that $M_d = f^{-1}(q)$, where q is the point of S_n with the coordinates $q = (1, 0, \cdots, 0)$. Furthermore, the manifold M'_d, which is the locus of the endpoint of the vector $\epsilon \boldsymbol{v}_1(x)$ as x runs over M_d (ϵ fixed, a small positive real number) is also the reverse image $M'_d = f^{-1}(q')$ of some point $q' \in S_n$ (precisely, q' is the point $q' = (1 - 2\epsilon^2, 2\epsilon(1-\epsilon^2)^{\frac{1}{2}}, 0, \cdots, 0)$, if the mapping $r\colon B_n \to S_n$ is indeed chosen as in (1.1)).

By the original definition of Hopf's invariant, one has

$$L(M_d, M'_d) = h(f) \mod 2, \tag{8.1}$$

where $L(\ ,\ )$ denotes the looping coefficient in E_{2d+1}.

Considering E_{2n-1} $(= E_{2d+1})$ as the linear subspace of E_{2n} defined by $y_{2n} = 0$, $y_1, y_2, \cdots, y_{2n-1}$ being coordinates in E_{2n-1}, we show first that there exists in E_{2n} an immersed (not necessarily orientable) manifold X_n, the regular boundary (mod 2) of which is the given manifold M_d imbedded in E_{2n-1}.

The existence of an abstract C^∞-manifold V_n with regular boundary (mod 2) diffeomorphic to M_d follows, by a theorem of R. Thom (see [14], Théorème IV. 10), from the fact that all Stiefel-Whitney numbers of M_d vanish mod 2 (this because the normal bundle over M_d is trivial). We want to prove the existence of an immersion $i\colon V_n \to E_{2n}$, such that $i|\partial V_n$ be the given imbedding $f\colon M_d \to E_{2n-1}$.

By Theorem 1 in Whitney's paper [15], there exists an analytic manifold A_n in euclidean $(2n+1)$-space E, which is C^2-homeomorphic to V_n. Map A_n into E_{2n} by F_0 defined as follows: $F_0|\partial A_n$ is the given imbedding f of $M = \partial A_n$ into E_{2n-1}. Consider a neighborhood $N \approx \partial A_n \times I$ of ∂A_n in A_n and represent points $u \in N$ by pairs (x, t), where $x \in \partial A_n$ and $0 \leqq t \leqq 1$. Define $F_0(u) = (fx, t) =$ the point of E_{2n} with $(2n-1)$ first coordinates coinciding with those of fx and the $2n$-th coordinate of which is t. Extend F_0 over A_n, such that $F_0(A - N) \subset \{y_{2n} > 1\}$. Let N_1 be the subset of N characterized by $0 \leqq t \leqq \frac{1}{3}$. Extending again F_0 over E now, we use Weier-

strass' approximation theorem to get a C^2 map $F_1\colon A_n \to E_{2n}$, such that $F_1\overline{(A-N_1)} \cap E_{2n-1} = 0$. Let ω^0 be a real valued C^2 function on A_n, such that $\omega^0 = 1$ in $A - N$, $\omega^0 = 1$ in N for $t \geqq \frac{2}{3}$, $0 \leqq \omega^0 \leqq 1$ for $\frac{1}{3} \leqq t \leqq \frac{2}{3}$, $\omega^0 = 0$ for $t \leqq \frac{1}{3}$. Take, for instance, $\omega^0(x,t) = \frac{1}{4}(1 + \cos(3\pi t))^2$ for (x,t) in N with $\frac{1}{3} \leqq t \leqq \frac{2}{3}$. For $u \in N$, $F_1(u)$ has the form $F_1(u) = F_0(u) + \xi(u)$. Define $F\colon A_n \to E_{2n}$ by $F|A - N = F_1|A - N$ and $F(u) = F_0(u) + \omega^0(u)\xi(u)$. Then F is C^2 and its restriction over ∂A_n is f. Moreover, $F|N_1 = F_0|N_1$, and thus F is completely regular in N_1 since F_0 is.

Now, by a theorem of H. Whitney (see [15], Theorem 2, assertions (a) and (b)), we can approximate F, together with its first derivatives by a completely regular immersion $j\colon A_n \to E_{\ n}$. Let N_2 be the subset of N_1 defined by $0 \leqq t \leqq \frac{1}{9}$ and define the C^2 real-valued function ω^1 over A_n by $\omega^1(x) = 1$ for $x \in A - N_1$, $\omega^1(u) = 1$ for $\frac{2}{9} \leqq t \leqq \frac{1}{3}$, $\omega^1(u) = \frac{1}{4}\cdot(1 + \cos(9\pi t))^2$ for $\frac{1}{9} \leqq t \leqq \frac{2}{9}$ and $\omega^1(u) = 0$ for $u \in N_2$. Substitute for j the immersion $i\colon A_n \to E_{2n}$ defined by $i|A - N_1 = j|A - N_1$. In N_1, j takes the form $j(u) = F(u) + \eta(u)$. Define $i(u)$ by $i(u) = F(u) + \omega^1(u)\eta(u)$. Since we may take $\eta(u)$ together with its first derivatives arbitrarily small and the derivatives of $\omega^1(u)$ are all zero except the derivative with respect to t which is $\leqq 20$, it follows that, for suitably chosen j, the map i will be completely regular ($F|N_1 = F_0|N_1$ is completely regular). Moreover, $i|N_2 = F_0|N_2$, and therefore *the normal vector to M_d and tangent to $i(A_n)$ is the constant vector $\boldsymbol{v}_0$, normal to E_{2n-1} in E_{2n}.* Finally, we can obviously manage that $i(A - \partial A)$ has no common point with E_{2n-1}. Denote $i(A_n)$ by X_n.

We replace now the given field $\boldsymbol{F}_n$ of n-frames $\boldsymbol{v}_1(x), \cdots, \boldsymbol{v}_n(x)$ normal to M_d in E_{2n-1} by the $(n+1)$-field $\boldsymbol{F}_{n+1}$ of $(n+1)$-frames $\boldsymbol{v}_0(x), \boldsymbol{v}_1(x), \cdots, \boldsymbol{v}_n(x)$ consisting of the (constant) vector $\boldsymbol{v}_0(x)$ tangent to X_n and normal to M_d at $x \in M_d$, followed by the vectors of the field $\boldsymbol{F}_n$.

The homotopy class of the map $\omega'\colon M_d \to V_{2n,n+1}$ induced by the $(n+1)$-field $\boldsymbol{F}_{n+1}$ is represented by the same remainder mod 2, i. e., c, as the homotopy class of the map $\omega\colon M_d \to V_{2n-1,n}$ induced by the given field $\boldsymbol{F}_n$ (*we have assumed* d *odd*, which is no loss of generality in view of $\gamma = h = 0$ for d even by [8], Théorème IV. *Thus $n = d + 1$ is even, $n = 2s$*).

On the other hand, using certain vectors of $\boldsymbol{F}_{n+1}$ as a cross-section over $M_d \subset X_n$, we can define relative characteristic Stiefel-Whitney classes as follows: the n last vectors of $\boldsymbol{F}_{n+1}$ provide a cross-section over M_d in the principal normal bundle over X_n and lead in every positive dimension to a class $\bar{W}_R{}^i \bmod M_d$ $(1 \leqq i \leqq n)$. The vector field v_0 leads to the characteristic class $W_R{}^n$ of the tangent bundle to X_n. Using finally the vectors

of F_{n+1} altogether, we can define the relative n-dimensional class $S_R{}^n$ of the Whitney sum: tangent $\oplus$ normal bundles over X_n. According to Whitney duality for relative characteristic classes (Theorem (4.1)), we have[6]

$$(8.2)\qquad S_R{}^n = W_R{}^n + W^{n-1}\cdot \bar{W}_R{}^1 + \cdots + W^1\cdot \bar{W}_R{}^{n-1} + \bar{W}_R{}^n.$$

The remaining part of the proof of Lemma (1.2) consists in interpreting the several terms occurring in formula (8.2).

First,

$$(8.3)\qquad S_R{}^n = c\cdot A_R{}^n,$$

where $A_R{}^n$ denotes the generator of $H^n(X, M; \mathbf{Z}_2)$. The class $S_R{}^n$ is indeed the obstruction class to the extension of F_{n+1} over X_n under the only condition that it keeps being an $(n+1)$-framefield in E_{2n}. The formula (8.3) follows then from the fact that $V_{2n,n+1}$ is $(n-2)$-connected.

Since $W_R{}^n$ is the obstruction cohomology class to the extension of the vector $v_0(x)$ as a *tangent* vector field over X, we have by [2] (Satz I, page 549)

$$(8.4)\qquad W_R{}^n = \chi(X)\cdot A_R{}^n.$$

It has been proved in Lemma (7.14) that

$$(8.5)\qquad rA_R{}^n = W^{n-1}\cdot \bar{W}_R{}^1 + \cdots + W^1\cdot \bar{W}_R{}^{n-1}.$$

Let us prove now that

$$(8.6)\qquad \bar{W}_R{}^n = h(f)A_R{}^n.$$

Indeed, $\bar{W}_R{}^n$ is the obstruction class to the extension over X_n of the vector $v_1(x)$ of the field F_{n+1} as a *unit* vector field *normal* to X_n in E_{2n}. The extension is possible over the $(n-1)$-skeleton (a triangulation of X is taken, such that M is a subcomplex). Suppose this extension has been constructed. Extend then $v_1(x)$ in the interior of the n-simplexes of X_n as a *normal*

[6] Notice that we could prove (8.2) as follows: The restriction over M of the normal principal bundle $\mathfrak{N}$ over X is trivial. Let $K = X \cup C$ be obtained by attaching to X the cone over M. Then $\mathfrak{N}$ can be extended over K (i.e., the map $N: X \to B_{SO(n)}$ inducing $\mathfrak{N}$ can be extended to a map $K \to B_{SO(n)}$). K being realized in E_{2n}, define $T: K \to B_{SO(n)}$ by $T(x) = n$-plane orthogonal to $N(x)$ for every $x \in K$. This defines a bundle $\mathfrak{T}$ over K. Let w be the total Stiefel-Whitney class of $\mathfrak{T}$, $\bar{w}$ the total class of $\mathfrak{N}$. By excision, one has $H^+(K) \cong H^+(K, C) \cong H^+(X, M)$, where H^+ is the ring consisting of the elements of positive dimensions in the cohomology ring. It is not difficult to see that these isomorphisms send w^n into $S_R{}^n + W_R{}^n$ and w^i, $\bar{w}^j$ for $1 \leq i \leq n$, $1 \leq j \leq n$, into $W_R{}^i$, $\bar{W}_R{}^j$ respectively. Ordinary Whitney duality $w \cdot \bar{w} = 0$ goes thus over into formula (8.2) in particular. However, this raisonning does not apply in the more general situation considered in Part I.

vector of length $\leqq 1$. This is always possible if we allow $\boldsymbol{v}_1(x)$ to have length 0 in some interior points of some n-simplexes of X_n. Consider the locus X' of the endpoint of the vector $\epsilon\boldsymbol{v}_1(x)$ so extended (where $\epsilon > 0$ is smaller than the radius of a tubular neighborhood of X_n in E_{2n}; see [14], page 27). Denote by $V = X + \bar{X}$ the closed manifold obtained by adding to X its mirror image $\bar{X}$ with respect to the hyperplane E_{2n-1} in E_{2n}. We may choose the extension $\boldsymbol{v}_1(x)$, $x \in X_n$, such that X' and V are in general position in E_{2n}, so that we may determine the intersection coefficient mod 2 of V and X' by simply counting the intersection points. We assume all intersection points to be simple. We have then two kinds of intersection points of X' with V: Those arising from the impossibility of extending $\boldsymbol{v}_1(x)$ in the interior of an n-simplex of X_n as a normal unit vector (let their number be I), and those which arise from self-intersection points of X. It is clear that intersection points of the last category can occur only in *pairs,* thus their number is $2N$, where N is some integer.

By definition of $\bar{W}_R{}^n$, we have $\bar{W}_R{}^n = IA_R{}^n$ modulo 2. Furthermore, $I = I + 2N = S(V, X') = L'(V, M')$ modulo 2, where $L'(\ ,\ )$ denotes the looping coefficient in E_{2n} and M', as before, the locus of the endpoint of $\epsilon\boldsymbol{v}_1(x)$ as x runs over M_d ($\partial X' = M' \bmod 2$). Since $V \cap E_{2n-1} = M_d$, we have $L'(V, M') = L(M, M')$, where $L(\ ,\ )$ is again the looping coefficient in E_{2n-1}. Therefore, $I = L(M, M') = h(f) \bmod 2$. In other words, $\bar{W}_R{}^n = h(f) A_R{}^n$.

Using the formulae (8.3), (8.4), (8.5) and (8.6), the formula (8.2) translates into

$$c = \chi(X) + r + h(f) \mod 2. \tag{8.7}$$

It has been proved in [8], that

$$\chi^*(M) = \chi(X) + \rho \mod 2, \tag{8.8}$$

where ρ is the rank of the bilinear form $S(x, y)$ defined by the intersection coefficient in $H_s(X_n; \mathbf{Z}_2)$, $n = 2s$ (i. e. $S(x, y)$ is the intersection coefficient of the classes $x, y \in H_s(X_n; \mathbf{Z}_2)$). It is not difficult to see that $\rho = r$. We prove, however,

$$\chi^*(M) = \chi(X) + r \mod 2 \tag{8.9}$$

more simply by considering the exact cohomology sequence of the pair (X, M), i. e.,

$$H^s(X) \xleftarrow{h^*} H^s(X, M) \xleftarrow{\delta} H^{s-1}(M) \leftarrow \cdots - H^0(M)$$
$$\xleftarrow{i^*} H^0(X) \xleftarrow{h^*} H^0(X, M) \leftarrow 0.$$

Using the completely orthogonal pairing of $H^s(X)$ with $H^s(X,M)$ to $H^n(X,M)$, it follows from $h^*C\cdot Z=C\cdot Z$, $C,Z\in H^s(X,M)$ that the kernel of h^* in $H^s(X,M)$ consists exactly of those elements C for which $C\cdot Z=0$ for every $Z\in H^s(X,M)$. Thus, we have $r=$ (rank of $H^s(X,M)$) — (rank kernel h^*). Using the exactness of the above sequence, we obtain

$$r=p_s(X,M)-p_{s-1}(M)+p_{s-1}(X)-p_{s-1}(X,M)+\cdots$$
$$+(-1)^s p_0(M)+(-1)^{s-1}p_0(X)+(-1)^s p_0(X,M),$$

where $p_q(X,M)$, $p_q(M)$, $p_q(X)$ denote the ranks of $H^q(X,M)$, $H^q(M)$, $H^q(X)$ respectively.

Replacing in this formula $p_q(X,M)$ by $p_{n-q}(X)$ according to relative Lefschetz-Poincaré duality in $X \bmod M$, we obtain

$$r=(-1)^s\chi^*(M)+(-1)^{s-1}\sum_{0\leqq i\leqq s-1}(-1)^i p_i(X)+(-1)^s\sum_{s\leqq i\leqq n}(-1)^i p_i(X).$$

In other words,

$$\text{(8.10)}\quad \chi^*(M)=\{\sum_{0\leqq i\leqq s-1}(-1)^i p_i(X)-\sum_{s\leqq i\leqq n}(-1)^i p_i(X)\}+(-1)^s\cdot r,$$

from which formula (8.9) follows by reduction modulo 2.

Now, since by definition $\gamma(f)=c-\chi^*(M)$, formulae (8.7) and (8.9) complete the proof of the Lemma (1.2): $\gamma(f)=h(f)$.

9. Consequences of the Lemma (1.2). Using the Lemma (1.2) we may improve the generalized Curvatura Integra theorem. We obtain first from $h(f)=0$ if $n\leqq d$ (because then $Sq^{d+1}(u^n)=0$) the

THEOREM (9.1). *Let the closed differentiable manifold M_d be regularly imbedded in E_{d+n} with a field of normal n-frames $\mathbf{F}_n$ and $n\leqq d$, then the corresponding curvatura integra c does not depend on the imbedding nor on the n-field and is given by $c=\chi^*(M_d)$.*

Indeed, $c-\chi^*(M)=\gamma(f)$, where f is a map $S_{d+n}\to S_n$ as constructed in Section 1, (b). Since $\gamma(f)=h(f)=0$ because $n\leqq d$, the theorem follows.

Suppose that the integer d is such that every element in $\pi_{2d+n}(S_{d+1})$ has even (or zero) Hopf's invariant. Then every element in $\pi_{d+n}(S_n)$ with arbitrary n has zero generalized Hopf's invariant h. In this case we may therefore omit the restriction on n in the above theorem and obtain the

THEOREM (9.2). *Let the closed differentiable manifold M_d be regularly imbedded in any euclidean space E_{d+n} with a field of normal n-frames*

$\boldsymbol{F}_n$; *if d is such that there is no element of odd Hopf's invariant in* $\pi_{2d+1}(S_{d+1})$, *then the curvatura integra c corresponding to the imbedding and to* $\boldsymbol{F}_n$ *depends in fact only on M_d and is given by* $c = \chi^*(M_d)$.

We may replace in this theorem the assumption of an *imbedding* of M_d by the weaker one of an *immersion* (with self-intersections allowed). Indeed, if M_d is immersed in E_{d+n} with a field of normal n-frames $\boldsymbol{F}_n$ defining a curvatura integra c, we may imbed E_{d+n} as a linear subspace in E_{d+N} $(n \leqq N)$, with a field of normal N-frames $\boldsymbol{F}_N$ which consists of the n vectors of $\boldsymbol{F}_n$ followed by $(N-n)$ constant mutually orthogonal unit vectors normal to E_{d+n} in E_{d+N}. The curvatura integra corresponding to the new field $\boldsymbol{F}_N$ is again c and if N has been chosen sufficiently large $(d+1 \leqq N)$, a slight deformation in E_{d+N} of the immersion $M_d \rightarrow E_{d+N}$ will provide an *imbedding* of M_d into E_{d+N} with a field of normal N-frames obtained by continuous deformation from $\boldsymbol{F}_N$ (apply Theorem 2, case (c) of H. Whitney's paper [15] to obtain the imbedding and the covering homotopy theorem to obtain the desired field). The curvatura integra c still belongs to the new situation to which now Theorem (9.2) applies (d has not been changed). Thus we obtain the

THEOREM (9.2*). *The thorem* (9.2) *is still valid replacing the assumption of an "imbedding" of M_d by the assumption of an "immersion."*

Of special interest may be the case $n = 1$, originally considered by Hopf for d even. If d is odd and the manifold M_d assumed to be imbedded in E_{d+1}, then the curvatura integra c (the degree of the Gauss mapping $M_d \rightarrow S_d$, in this case) is modulo 2 equal to $\chi^*(M_d)$. This was proved in [8] and also by J. Milnor, using a simpler method especially adapted to this special case, in [9]. As a corollary to Theorem (9.2*), we obtain the following improvement:

COROLLARY. *Let d be an integer such that there is no element of* $\pi_{2d+1}(S_{d+1})$ *with odd Hopf's invariant. Let M_d be a closed orientable hypersurface in E_{d+1} with self-intersections allowed, and let c be the degree of the Gauss mapping $M_d \rightarrow S_d$. Then $c = \chi^*(M)$ modulo 2.*

Let us remark that the condition on d in the Theorem (9.2) (no map with odd Hopf's invariant in $\pi_{2d+1}(S_{d+1})$) is known to be satisfied at least for $d \neq 2^a - 1$, according to J. Adem [1] (explicit proofs in H. Cartan [7]).*

---

* Added in proof: The details of Adem's proof have been recently published in *Algebraic Geometry and Topology*, Princeton University Press, 1957.

According to H. Toda it is satisfied for $d=15$ too (As is well known, H. Hopf has proved that the condition is not satisfied for $d=1,3$ and 7).

Along the same lines, we obtain an extension of a part of N. Steenrod-J. H. C. Whitehead's theorem [13] according to which the d-sphere S_d cannot be parallelizable if $d \neq 2^a - 1$.

THEOREM (9.3). *Any manifold M_d with odd semi-characteristic $\chi^*(M)$ is not parallelizable if $d \neq 2^a - 1$, (in fact, if d is such that every element in $\pi_{2d+1}(S_{d+1})$ has even Hopf's invariant).*

Proof. Suppose M_d is parallelizable. It is known that any regular imbedding of M_d in E_{2d+1} admits in this case a field of normal $(d+1)$-frames $\boldsymbol{F}_{d+1}$, the corresponding curvatura integra c being zero (see [8], Section 8). Therefore, the semi-characteristic being odd, the corresponding value of γ is $\gamma = c - \chi^*(M) = 1$. Since by Lemma (1.2), γ is the Hopf's invariant mod 2 of some map $S_{2d+1} \to S_{d+1}$, the dimension d must be of the form $d = 2^a - 1$.

I do not know if there are examples of manifolds M_d with odd semi-characteristic, carrying 2^k fields of mutually orthogonal unit vectors (k being defined by $d+1 = 2^k(2r+1)$).

It was proved in [8] (Corollaire au théorème VIII, § 8), that the real projective space P_d may be immersed into a euclidean space E_{d+n} $(n \geqq d+1)$ with a field of normal n-frames if and only if it is parallelizable. The Lemma (1.2) provides a similar (perhaps weaker) statement for a larger class of manifolds:

THEOREM (9.4). *Suppose that the manifold M_d admits a $2m$-fold covering manifold $\bar{M}_d$ with odd semi-characteristic: $\chi^*(\bar{M}_d) = 1 \bmod 2$. Then the manifold M_d cannot be immersed into any euclidean space E_{d+n} with a field of normal n-frames unless there is in $\pi_{2d+1}(S_{d+1})$ some element of odd Hopf's invariant.*

Proof. From an immersion $g\colon M_d \to E_{d+n}$ of M_d into some euclidean $(d+n)$-space E_{d+n} with a field $\boldsymbol{F}_n$ of normal n-frames, we obtain by composition with the covering map $p\colon \bar{M}_d \to M_d$ a regular immersion $\bar{g}\colon M_d \to E_{d+n}$ with a field of normal n-frames $\boldsymbol{F}_n$. Let us call $\bar{c}$ the curvatura integra of $\bar{M}_d$ corresponding to $\bar{\boldsymbol{F}}_n$. Because $\bar{g}(\bar{M}_d)$ is homologous to $2m \cdot g(M_d)$ in $V_{d+n,n}$, $\bar{g}$ denoting the composition $\bar{g} = g \circ p$, we have $\bar{c} = 2m \cdot c = 0 \pmod 2$.

Therefore, $\bar{c} - \chi^*(\bar{M}_d) \neq 0$, and because of Theorem (9.2*), $\pi_{2d+1}(S_{d+1})$ must contain some element of odd Hopf's invariant. This completes the proof of Theorem (9.4).

It is known that if S_d is parallelizable, then any manifold M_d which may be immersed in some euclidean space E_{d+n} with a normal n-frame is also parallelizable (see [8], Théorème VIII). In [9], J. Milnor formulates the conjectures that if a parallelizable manifold M_d may be immersed in E_{d+1} in such a way that the Gauss degree be 1, then S_d should be parallelizable.

From Lemma (1.2) follows the

THEOREM (9.5). *If a parallelizable manifold M_d may be immersed in some euclidean space E_{d+n} with a field of normal n-frames inducing an odd curvatura integra, then there is in $\pi_{2d+1}(S_{d+1})$ some element of odd Hopf's invariant.*

Proof. Assuming $n \geqq d+1$, and using the covering homotopy theorem, we may construct on M_d a field G_n of normal n-frames in E_{d+n}, such that the induced curvatura integra is zero (see [8], § 8). Thus the given field and G_n induce different curvatura integra. By Theorem (9.2*), this implies the existence in $\pi_{2d+1}(S_{d+1})$ of an element of odd Hopf's invariant (the question whether this is possible without S_d being parallelizable is unsolved as is well known).

10. Remark. It should be noticed that if one is not interested in the value of the curvatura integra c but only in the fact that c does not depend on the imbedding nor on the normal field (as in Theorem 9.5), then a simpler proof may be given in the case d odd. Let us sketch a "direct" proof of the following weaker form of Theorem (9.2) in this case:

THEOREM (10.1). *If the odd integer d is such that each element of $\pi_{2d+1}(S_{d+1})$ has even Hopf's invariant, then the curvatura integra of any closed differentiable manifold M_d regularly immersed into E_{d+n} with a field of normal n-frames depends only on M_d.*

Proof. Let $M_d{}^i = T^i(M_d)$, $i = 1, 2$ be two regular immersions of the given manifold M_d of dimension d into euclidean spaces E_{d+n_i}. Let $F^i{}_{n_i}$ be two fields $(i = 1, 2)$ of normal n_i-frames on $M_d{}^i$ respectively, inducing maps $\phi^i\colon M_d \to V_{d+n_i, n_i}$ the classes of which are represented by c^i. In order to prove $c^1 = c^2$, let $M_d{}^0 = T^0(M_d)$ be an arbitrary regular *imbedding* of M_d into E_{2d+1} (euclidean $(2d+1)$-space). Consider $E_{d+n_1}, E_{2d+1}, E_{d+n_2}$ as linear subspaces of the euclidean space $E_N = E_{d+n_1} \times E_{2d+1} \times E_{d+n_2}$. Consider on $M_d{}^i$ the fields $F^i{}_{N-d}$ of $(N-d)$-frames normal to $M_d{}^i$, consisting of the vectors of $F^i{}_{n_i}$ followed by $N-d-n_i$ constant unit vectors (mutually

orthogonal) normal to E_{d+n_i} in E_N. Each immersion $T^i(M_d)$, $i=1,2$, is isotopic in E_N with $T^0(M_d)$. By the covering homotopy theorem, we obtain on $M_d{}^0$ two fields of normal $(N-d)$-frames in E_N, which we denote again by $\boldsymbol{F}^i{}_{N-d}$. The curvatura integra corresponding to the mapping $M_d \rightarrow V_{N,N-d}$ induced by the new field $\boldsymbol{F}^i{}_{N-d}$ on $M_d{}^0$ is c^i.

It is easily seen that $\boldsymbol{F}^i{}_{N-d}$ may be continuously deformed (keeping the $(N-d)$-frames of $\boldsymbol{F}^i{}_{N-d}$ normal to $M_d{}^0$ during the deformation), in such a way that the first $(d+1)$ vectors become vectors in E_{2d+1} and the last $N-(2d+1)$ be constant and normal to E_{2d+1}. In other words, we have obtained two fields $\boldsymbol{F}^i{}_{d+1}$, $i=1,2$, of $(d+1)$-frames normal to $M_d{}^0$ in E_{2d+1}. Moreover, the curvatura integra corresponding to the map $M_d \rightarrow V_{2d+1,d+1}$ induced by $\boldsymbol{F}^i{}_{d+1}$ (as field over $M_d{}^0$ in E_{2d+1}) is equal to the given c^i which we started from.

Recall that $T^0(M_d)=M_d{}^0$ has been assumed to be an *imbeddding* into E_{2d+1}. According to Section 1 (b), it thus corresponds to $\boldsymbol{F}^1{}_{d+1}$ and $\boldsymbol{F}^2{}_{d+1}$ sphere maps f^1, f^2 of S_{2d+1} into S_{d+1}. By assumption, these maps have the same Hopf's invariant mod 2: $h(f^1)=h(f^2) \bmod 2$. It is not difficult to see, that we may change one of the fields, $\boldsymbol{F}^1{}_{d+1}$ say, without changing c^1 (which is only defined mod 2 because d has been assumed to be odd) in such a way that $h(f^1)=h(f^2)$ as integers. Assume that such a change has been achieved. The Hopf's invariant of f^i is the looping coefficient in E_{2d+1} of $M_d{}^0$ with the locus, $V_d{}^i$ say, of the endpoint of the first vector of the field $\boldsymbol{F}^i{}_{d+1}$. These looping coefficients being equal, it is possible using again the covering homotopy theorem, to deform $\boldsymbol{F}^1{}_{d+1}$ and $\boldsymbol{F}^2{}_{d+1}$ continuously (keeping their unit vectors mutually orthogonal and normal to $M_d{}^0$) in such a way that after the deformation their first vectors coincide. Such a deformation does not change c^1 or c^2.

Let us denote by $\{\boldsymbol{v}, \boldsymbol{v}_1{}^i, \boldsymbol{v}_2{}^i, \cdots, \boldsymbol{v}_d{}^i\}$ the vectors of $\boldsymbol{F}^i{}_{d+1}$ (after deformation) and by $\theta^i\colon M_d \rightarrow V_{2d+1,d+1}$ the induced mappings, the classes of which are represented by c^i $(i=1,2)$.

In order to prove $c^1=c^2$, let us first assume that M_d is a parallelizable manifold and let $\boldsymbol{t}_1, \boldsymbol{t}_2, \cdots, \boldsymbol{t}_d$ be a d-field of (mutually orthogonal unit) tangent vectors on $M_d{}^0$. Then each θ^i is homotopic to the map $\theta\colon M_d \rightarrow V_{2d+1,d+1}$ defined by $\theta(x)=\{\boldsymbol{v}(x), \boldsymbol{t}_1(x), \boldsymbol{t}_2(x), \cdots, \boldsymbol{t}_d(x)\}$. The desired homotopy is given by

$$\theta^i{}_s(x)=\{\boldsymbol{v}(x), \boldsymbol{v}_1{}^i(x)\cos(\tfrac{1}{2}\pi s)+\boldsymbol{t}_1(x)\sin(\tfrac{1}{2}\pi s), \cdots, \\ \boldsymbol{v}_d{}^i(x)\cos(\tfrac{1}{2}\pi s)+\boldsymbol{t}_d\sin(\tfrac{1}{2}\pi s)\},$$

where $0 \leqq s \leqq 1$. Therefore, if M_d is parallelizable, θ^1 and θ^2 are homotopic and $c^1 = c^2$.

In general, M_d will not be parallelizable. However, if c^1 and c^2 were different, then one of them would be zero (it follows from d being odd, that c^1 and c^2 are remainders mod 2). If $c^i = 0$, $i = 1$ or 2, the corresponding map θ^i is homotopic to zero. By a reasoning of [7] (§ 8), it follows that M_d would be parallelizable.

This completes the "direct" proof of Theorem (10.1).

Appendix.

11. Relative Chern and Pontryagin characteristic classes.

11.a. *Relative Chern classes.*

In this section the coefficients are the *integers*.

Let $\mathfrak{B} = (E_{\boldsymbol{U}(n)}, p, B_{\boldsymbol{U}(n)}, \boldsymbol{U}(n))$ be the classifying bundle for the unitary group of n variables $\boldsymbol{U}(n)$. Suppose that a cross-section θ^r over a closed subset A of $B_{\boldsymbol{U}(n)}$ is given in the associated bundle $\mathfrak{B}^r$ with fibre $W_{n,n-r}$ (the complex Stiefel manifold of $n-r$ complex vectors in C_n).

For $q \geqq r$, the relative Chern class $C_R^{q+1} \in H^{2(q+1)}(B_{\boldsymbol{U}(n)}, A; \boldsymbol{Z})$ corresponding to the cross-section θ^r will be defined by the properties

(11.1) $a^* C_R^{q+1} = C^{q+1}$, the ordinary (absolute) Chern class, a^* being the homomorphism $H^*(B_{\boldsymbol{U}(n)}, A, \boldsymbol{Z}) \to H^*(B_{\boldsymbol{U}(n)}, \boldsymbol{Z})$ induced by the inclusion $a\colon (B_{\boldsymbol{U}(n)}, 0) \to (B_{\boldsymbol{U}(n)}, A)$,

(11.2) $\rho^*_{q,n} C_R^{q+1} = 0$, where $\rho^*_{q,n}\colon H^*(B_{\boldsymbol{U}(n)}, A) \to H^*(B_{\boldsymbol{U}(q)}, \theta^q A)$ is induced by the Borel map $\rho(\boldsymbol{U}(q), \boldsymbol{U}(n))$.

We consider the diagram

$$\begin{tikzcd}
\cdots \arrow[r] & H^q(B_{\boldsymbol{U}(q)}) \arrow[r] & H^q(\theta^q A) \arrow[r, "\delta"] \arrow[d, "\theta^*"] & H^{q+1}(B_{\boldsymbol{U}(q)}, \theta^q A) \arrow[r, "\bar{a}^*"] & H^{q+1}(B_{\boldsymbol{U}(q)}) \arrow[r] & \cdots \\
\cdots \arrow[r] & H^q(B_{\boldsymbol{U}(n)}) \arrow[r] \arrow[u, "\alpha^*"] & H^q(A) \arrow[r, "\delta"] & H^{q+1}(B_{\boldsymbol{U}(n)}, A) \arrow[r, "a^*"] \arrow[u, "\rho^*_{q,n}"] & H^{q+1}(B_{\boldsymbol{U}(n)}) \arrow[r] \arrow[u, "\alpha^*"] & \cdots
\end{tikzcd}$$

where $B_{\boldsymbol{U}(q)}$ is the space of the bundle $\mathfrak{B}^q$ ($B_{\boldsymbol{U}(q)}$ is a classifying space for $\boldsymbol{U}(q)$, thus the notation) and θ^q is the cross-section over A in $\mathfrak{B}^q$ induced by θ^r $(q \geqq r)$.

By considerations similar to those made for the orthogonal group, it is easily seen that α^* is an epimorphism in every dimension and a monomor-

phism in dimensions not exceeding q. It follows, using exactness and commutativity in the diagram, that if $\rho^*_{q,n}z = 0$ and $a^*z = 0$ for some $z \in H^*(B_{U(n)}, A)$ then z must be zero.

The existence of at least one cohomology class with properties (11.1) and (11.2) is seen as follows: The restriction of C^{q+1} to A is zero because of the assumed existence of a cross-section over A in $\mathfrak{B}^r$. Let $C^{q+1} = \underset{\cdot}{a}^*x$. Because $0 = \alpha^*C^{q+1} = \alpha^*a^*x = \bar{a}^*\rho^*_{q,n}x$, we have $\rho^*_{q,n}x = \delta\theta^{*-1}y$ for some $y \in H^q(A)$. The class $x - \delta y$ has the properties (11.1) and 11.2).

This proves that properties (11.1) and (11.2) indeed define the relative Chern classes uniquely for the classifying $\boldsymbol{U}(n)$-bundle.

We consider now the more general situation of a $\boldsymbol{U}(n)$-bundle over some compact finite dimensional space X induced by some map $g\colon X \to B_{U(n)}$. Let $(E^r, \pi, X, W_{n,n-r})$ be the associated bundle with fibre $W_{n,n-r}$ and assume a cross-section $\theta^r\colon A \to E^r$ to be given over the closed subset A of X. We may assume $\dim B_{U(n)}$ arbitrarily high. Take $\dim B_{U(n)} \geqq 2 \dim X + 1$ and let $f\colon X \to B_{U(n)}$ be an *injective* map homotopic to g. The bundle induced by f is equivalent to the one induced by g; let us denote it again by $(E^r, \pi, X, W_{n,n-r})$.

Let S be a closed subset of $B_{U(n)}$ containing $f(A)$ and such that there exists a cross-section $\psi\colon S \to B_{U(r)}$ in $\mathfrak{B}^r$, with the property $\psi(a) = \bar{f}\theta(a)$ for every $a \in A$ ($\bar{f}\colon E^r \to B_{U(r)}$ is the bundle map covering $f\colon X \to B_{U(n)}$). Let c_R^{q+1} be the $(q+1)$-dimensional relative Chern class of the classifying bundle mod S obtained using the cross-section $\psi (q \geqq r)$. We shall prove the

LEMMA (11.3). *$f^*c_R^{q+1}$ depends only on the homotopy class of the map g inducing the given bundle and on the cross-section θ^r over A.*

Definition. $f^*c_R^{q+1} = C_R^{q+1} \in H^{2(q+1)}(X, A; \mathbf{Z})$ is the relative Chern class (of dimension $2(q+1)$, defined for $q \geqq r$) mod A of the bundle (E, π, X), corresponding to the cross-section θ^r.

In order to prove Lemma (11.3), we first notice that $f^*c_R^{q+1}$ does not depend on S. Indeed, let $i^*\colon H^*(B_{U(n)}, S) \to H^*(B_{U(n)}, f(A))$ be the homomorphism dual to the inclusion $i\colon (B_{U(n)}, f(A)) \to (B_{U(n)}, S)$. We prove that $i^*c_R^{q+1}$ is the relative Chern class mod $f(A)$ corresponding to the restriction ψ_A of ψ over $f(A)$. Consider the diagram

$$\begin{array}{ccccc} & & H^{2(q+1)}(B_{U(n)}, fA) & \xleftarrow{\;i^*\;} & H^{2(q+1)}(B_{U(r)}, \psi S) \\ & & \Big\uparrow \rho^*_{r,n} & & \Big\uparrow \rho'^*_{r,n} \\ H^{2(q+1)}(B_{U(n)}) & \xleftarrow{\;a^*\;} & H^{2(q+1)}(B_{U(r)}, \psi f A) & \xleftarrow{\;\bar{i}^*\;} & H^{2(q+1)}(B_{U(n)}, S). \end{array}$$

By commutativity, the relations $\rho^*_{r,n}(i^*c_R{}^{q+1})=0$ and $a^*(i^*c_R{}^{q+1})=C^{q+1}$ (showing, by properties (11.1) and (11.2), that $i^*c_R{}^{q+1}$ is indeed the relative Chern class mod fA corresponding to ψ_A) follow from the corresponding relations $\rho'^*_{r,n}c_R{}^{q+1}=0$ and $a'^*c_R{}^{q+1}=c^{q+1}$ for $c_R{}^{q+1}$ (where $a'^*=a^*i^*$). Since $i^*c_R{}^{q+1}$ is independent of S, so is $f^*c_R{}^{q+1}$, since $f^*=f'^*i^*$, where f'^* is induced by $f'\colon(X,A)\to(B_{U(n)},f(A))$.

It remains to be proved that $f^*c_R{}^{q+1}$ does not depend on the choice of the injective map. Let f_1, f_2 be two injective maps $X\to B_{U(n)}$ homotopic to g. We may assume that $f_1A\cap f_2A=0$. Otherwise take an injective map $f_0\colon X\to B_{U(n)}$ such that $f_0X\cap f_1X=0$ and $f_0X\cap f_2X=0$ (such a map may be obtained taking $\dim B_{U(n)}\geqq 2\dim X+3$ if necessary) and apply the following proof to f_0 and f_1 first and again to f_0 and f_2. Let S be the union $f_1A\cup f_2A$. The cross-section ψ over S is given by $\psi(a_i)=\theta f_i^{-1}(a_i)$ for $a_i\in f_i(A)$. Denoting by c_{1R} and c_{2R} the relative (universal) Chern classes mod f_1A and f_2A respectively corresponding to the restriction of ψ over f_1A and f_2A, we have to prove $f_1{}^*c_{1R}=f_2{}^*c_{2R}$. By the above remark, we have $f_1{}^*c_{1R}=f_1{}^*c_R$, $f_2{}^*c_{2R}=f_2{}^*c_R$, where c_R is the $2(q+1)$-dimensional class mod S corresponding to ψ. The equality $f_1{}^*c_R=f_2{}^*c_R$ due to $f_1\simeq f_2$ completes the proof of the Lemma (11.3).

Remark. Similarly to the definition of the Stiefel-Whitney classes, the definition of the relative Chern classes could have been introduced in terms of obstructions to the extension of cross-sections originally given over a subcomplex of the base. The two definitions coincide on their common domain: $\boldsymbol{U}(n)$-bundles $\mathfrak{B}$ over a complex K modulo some subcomplex L. (Use S. Eilenberg's approximation theorem, l. c. section 5).

Naturality property.

LEMMA (11.4). *Let (E,π,X) be a $\boldsymbol{U}(n)$-bundle and let (E',π',X') be the $\boldsymbol{U}(n)$-bundle over X' induced by some map $g\colon X'\to X$. We denote by C'_R and C_R the $2(q+1)$-dimensional relative Chern classes of the bundles (E',π',X') and (E,π,X) respectively, modulo closed subsets $A'\subset X'$ and $A\subset X$ such that $g(A')\subset A$ and corresponding to cross-sections θ and θ' in the associated bundles with fibre $W_{n,n-r}$ $(r\leqq q)$, such that $g\theta'(a')=\theta g(a')$ for every $a'\in A'$. Then $C'_R=g^*(C_R)$.*

Proof. Let $f\colon X\to B_{U(n)}$ be an injective map inducing (E,π,X) and $f'\colon X'\to B_{U(n)}$ be an injective map inducing (E',π',X'). We may assume that $fg(A')$ and $f'(A')$ have no commoon point. Let S be the union $fg(A')\cup f'(A')$ and ψ the cross-section over S in the associated bundle with

fibre $W_{n,n-r}$ defined by $\psi f'(a') = \bar{f}'\theta'(a')$ and $\psi fg(a') = \bar{f}\theta g(a')$. Denoting again by c_R the $2(q+1)$-dimensional (universal) relative Chern class modulo S corresponding to ψ, we have $C'_R = f'^*(c_R) = g^*f^*(c_R) = g^*(C_R)$.

Whitney duality. Let $\mathfrak{B}_1$ and $\mathfrak{B}_2$ be two principal bundles with bundle groups $\boldsymbol{U}(n_1)$ and $\boldsymbol{U}(n_2)$ respectively, over the same base space X and let $\theta_1^{r_1}$ and $\theta_2^{r_2}$ be cross-sections over closed subsets A_1 and A_2 (of X) in the associated bundles $\mathfrak{B}_1^{r_1}$ and $\mathfrak{B}_2^{r_2}$ with fibres W_{n_1,n_1-r_1} and W_{n_2,n_2-r_2} respectively. $\theta_1^{r_1}$ and $\theta_2^{r_2}$ determine a cross-section θ^r $(r = r_1 + r_2)$ over $A = A_1 \cap A_2$ in the bundle $\mathfrak{B}^r$, with fibre $W_{n,n-r}$ $(n = n_1 + n_2)$ associated to the Whitney sum $\mathfrak{B} = \mathfrak{B}_1 \oplus \mathfrak{B}_2$.

Denote by C_{iR}^{q+1} the relative Chern class of $\mathfrak{B}_i$, $i = 1, 2$, defined for $q \geqq r_i$, and let C_R^{q+1} be the relative Chern class of $\mathfrak{B}$ defined for $q \geqq r$.

For the relative Chern classes, the Whitney duality takes the form

$$C_R^{q+1} = C_{1_R}^{q+1} + C_{1_R}^{q} \cdot C_2^{1} + \cdots + C_{2_R}^{q+1}. \tag{11.5}$$

In (11.5) some absolute Chern classes occur, however, again, since each product contains at least one relative class, it is itself a relative class.

The proof of formula (11.5) will be based on a theorem similar to Theorem (5.1) stated below (Theorem (11.6)).

Relative Chern classes as symmetric functions. Let $(E, p, X, \boldsymbol{U}(n))$ be a principal $\boldsymbol{U}(n)$-bundles over a compact finite dimensional space X. Consider the subgroup $\boldsymbol{Q}(n) = \boldsymbol{U}(1) \times \boldsymbol{U}(1) \times \cdots \times \boldsymbol{U}(1)$, n factors, of $\boldsymbol{U}(n)$ and the "space of flags" $\bar{X} = E/\boldsymbol{Q}(n)$ over X. We have a fibering $\rho\colon \bar{X} \to X$ induced by the projection $p\colon E \to X$ and the cross-section θ^r over $A \subset X$ may be used to construct a subset $\bar{A} \subset \bar{X}$ as follows. Let $E^r = E/\boldsymbol{U}(r)$ be the space of the bundle with fibre $W_{n,n-r}$ associated to $(E, p, X, \boldsymbol{U}(n))$. We have the diagram

$$\begin{tikzcd}
\bar{A} \arrow[r, "\bar{\theta}"] \arrow[d, "\rho_A"] & \bar{E}^r \arrow[r, "\omega"] \arrow[d, "\rho_E"] & \bar{X} \arrow[d, "\rho_X"] \\
A \arrow[r, "\theta"'] & E^r \arrow[r, "p"'] & X
\end{tikzcd}$$

where $\bar{E}^r$ is the space of flags over E^r (i.e. $E/\boldsymbol{Q}(r) = \bar{E}^r$) and $\rho_E\colon \bar{E}^r \to . E^r$ is induced by $\boldsymbol{Q}(r) \subset \boldsymbol{U}(r)$. The bundle $(\bar{A}, \rho_A, A, \boldsymbol{F}(r), \boldsymbol{U}(r))$ is induced by $\theta\colon A \to E^r$. The principal bundle $\mathfrak{E} = (E, \pi, \bar{X}, \boldsymbol{Q}(n))$ is the Whitney sum $\mathfrak{E} = \mathfrak{E}^1 \oplus \mathfrak{E}^2 \oplus \cdots \oplus \mathfrak{E}^n$ of n principal bundles $\mathfrak{E}^i$ with group $\boldsymbol{U}(1)$. Let $x_i \in H^2(\bar{X}; \boldsymbol{Z})$ be the 2-dimensional (only non-zero positive dimensional)

Chern class of $\mathfrak{E}^i$. For $r+1 \leqq i \leqq n$, the relative Chern class mod $\bar{A}$ of $\mathfrak{E}^i$ may be defined (using the cross-section over $\bar{A}$ in $\mathfrak{E}^i$ induced by θ^r). Let $x_{r+1}, x_{r+2}, \cdots, x_n$ mean the *relative* classes. Then the elementary symmetric function $S^{q+1}(x_1, \cdots, x_n)$ is a relative class for $q \geqq r$ (each product of $q+1$ distinct factors from the $x_1, \cdots, x_n$ must contain at least one of the variables $x_{r+1}, \cdots, x_n$). Let us denote by $S_R^{q+1}(x_1, \cdots, x_n)$ this relative class. We have the theorem

THEOREM (11.6). *Let $\rho_R^*: H^*(X, A; \mathbf{Z}) \to H^*(\bar{X}, \bar{A}; \mathbf{Z})$ be the homomorphism induced by $\rho_R: (\bar{X}, \bar{A}) \to (X, A)$. Then,*

(a) *ρ_R^* is a monomorphism,*

(b) *$\rho_R^* C_R^{q+1} = S_R^{q+1}(x_1, \cdots, x_n)$ for $q \geqq r$.*

The proof is very similar to the proof of Theorem (5.1). We are not going to enter into the details again. Along the same lines as for the proof of (5.1), we need the result

$$H^*(\boldsymbol{F}(n); \mathbf{Z}) = \mathbf{Z}[x_1, \cdots, x_n]/(S^+(x_1, \cdots, x_n)),$$

with $\boldsymbol{F}(n) = \boldsymbol{U}(n)/\boldsymbol{Q}(n)$, where $(S^+(x_1, \cdots, x_n))$ denotes the ideal generated in $\mathbf{Z}[x_1, \cdots, x_n]$ by the symmetric functions of positive degree (see [3], Proposition 31.1). The only points where the method of proof of (5.1) breaks down now are those where use was made of the fact that the coefficient ring for Stiefel-Whitney classes was a *field* $(\mathbf{Z}_2)$.

The assertion that the term E_2 in the spectral sequence of $\rho: \bar{X} \to X$ reduces to $H^*(X) \otimes H^*(\boldsymbol{F}(n))$ is, however, still true because $\boldsymbol{F}(n) = \boldsymbol{U}(n)/\boldsymbol{Q}(n)$ has no torsion.

The only non-trivial change is that it is no longer obvious that $H^*(X; \mathbf{Z})$ should be (at least additively) isomorphic to E_∞. This proves, however, to hold, as will be seen from the following argument I learned from A. Borel, similar to an argument by J. P. Serre ([10], Chap. III, 7, Prop. 9).

LEMMA (11.7). *Let $\sigma: H^*(F) \to H^*(E)$ be a right-inverse (additive) homomorphism to $i^*: H^*(E) \to H^*(F)$, where E is the space of a bundle (E, π, B, F) with fibre F and i the inclusion $i: F \to E$. Assume that either B or F has no torsion. Then $H^*(E)$ is isomorphic to $H^*(B) \otimes H^*(F)$. Moreover, the isomorphism preserves the product if σ does.* (In this lemma the domain of coefficients is any commutative ring with unit).

Proof. Let $\omega: H^*(B) \otimes H^*(F) \to H^*(E)$ be the linear map defined

by $\omega(x\otimes f)=\pi^*(x)\cdot\sigma(f)$. We are going to prove that ω is an isomorphism which preserves products if σ does.

We first notice that ω preserves the filtration: $H^*(B)\otimes H^*(F)$ being filtered by the ideals $A^q=\sum_{i\geqq q}H^i(B)\otimes H^*(F)$ and $H^*(E)$ by the ideals $J^q=\pi^*(\sum_{i\geqq q}H^i(B))\cdot H^*(E)$. We have for every $q,\omega(A^q)\subset J^q$. Therefore, ω induces an additive homomorphism $\bar{\omega}$ of the corresponding graded rings $\sum A^q/A^{q+1}$ into $\sum J^q/J^{q+1}$, where $\sum$ denotes the direct sum. It is well known and easily seen that ω is an isomorphism if $\bar{\omega}$ is.

By triviality of the spectral sequence of the fibering $p\colon E\to B$ (because $i^*\colon H^*(E)\to H^*(F)$ which admits a right inverse, is an epimorphism), it follows that $E_2=H^*(B,H^*(F))$ is isomorphic to the graded ring E_∞ associated to $H^*(E)$, i.e. $\sum J^q/J^{q+1}$. Let k be this isomorphism ($k^2{}_\infty$ with the notations of [3], § 1).

Since either B or F has no torsion by assumption, $E_2=H^*(B)\otimes H^*(F)$; identifying $\sum A^q/A^{q+1}$ with $H^*(B)\otimes H^*(F)$ in the natural manner, the lemma will be proved by showing that $\bar{\omega}\colon H^*(B)\otimes H^*(F)\to E_\infty$ is identical to the map $k\colon H^*(B)\otimes H^*(F)\to E_\infty$.

Let $x\otimes f\in H^*(B)\otimes H^*(F)$, where $x\in H^p(B)$, $f\in H^q(F)$. We have $\omega(x\otimes f)=\pi^*(x)\cdot\sigma(f)\in J^{p,q}=\pi^*(\sum_{i\geqq p}H^i(B))\cdot H^q(E)$, and $\bar{\omega}(x\otimes f)$ is the image of $\omega(x\otimes f)$ in $J^{p,q}/J^{p+1,q-1}=E_\infty{}^{p,q}$. By definition of the product in E_∞, we obviously have $\bar{\omega}(x\otimes f)=\bar{\omega}(x\otimes 1)\cdot\bar{\omega}(1\otimes f)$. Since k has also this property, it is sufficient to verify $\bar{\omega}=k$ on the elements of the form $x\otimes 1$ and $1\otimes f$.

$\bar{\omega}(x\otimes 1)=\pi^*(x)\in J^{p,0}$ $(J^{p+1,-1}=0)$. We have $k(x\otimes 1)=\pi^*(x)$. See [3], § 4, (b).

$\bar{\omega}(1\otimes f)=\overline{\sigma(f)}$, where $\overline{\sigma(f)}$ represents the image of $\sigma(f)\in J^{0,q}$ in $J^{0,q}/J^{1,q-1}$. By [3], § 4 (c), we have $k^{-1}(\overline{\sigma(f)})=1\otimes i^*\sigma(f)=1\otimes f$ since σ is a right inverse to i^*. Thus, again $\bar{\omega}(1\otimes f)=k(1\otimes f)$.

If σ preserves the product, then

$$\begin{aligned}\omega[(x\otimes f)\cdot(x'\otimes f')]&=(-1)^{qp'}\omega(xx'\otimes ff')=(-1)^{qp'}\pi^*(x\cdot x')\cdot\sigma(f\cdot f')\\&=(-1)^{qp'}\pi^*(x)\cdot\pi^*(x')\cdot\sigma(f)\cdot\sigma(f')=\pi^*(x)\cdot\sigma(f)\cdot\pi^*(x')\cdot\sigma(f')\\&=\omega(x\otimes f)\cdot\omega(x'\otimes f').\end{aligned}$$

In the situation of Theorem (11.6), the existence of the homomorphism σ is obvious (one has $H^*(\boldsymbol{F}(n);\boldsymbol{Z})=\boldsymbol{Z}[u_1,\cdots,u_n]/(S^+(u_1,\cdots,n_n))$, where $u_k=i^*x_k$; let $h_1=1,h_2,\cdots,h_t$ be a basis for the vector space

$H^*(\boldsymbol{F}(n);\boldsymbol{Z})$, the h_i being represented by polynomials $P_i(u_1,\cdots,u_n)$ in $u_1,\cdots,u_n$, and define $\sigma(h_i)$ as the polynomial in $x_1,\cdots,x_n$ obtained from P_i by the substitution $u_k \to x_k$; extend then σ by linearity). This completes the proof of Theorem (11.6) part (a).

The proof of part (b) is similar to the proof of (b) in Theorem (5.1). Use has to be made of the results of [6] (in particular Prop. 4.1) rather than from [4]. A slight change occurs at the end of the proof: we have to show that $\mu_R{}^*S_R{}^{q+1}(x_1,\cdots,x_n)=0$, where $\mu_R{}^*\colon H^*(B_{\boldsymbol{Q}(n)},\bar{A}) \to H^*(B_{\boldsymbol{Q}(r)},\bar{A}^1)$. The argument used for Stiefel-Whitney classes does not work now, because $\deg x_i = 2$ and $a^*\colon H^2(B_{\boldsymbol{Q}(r)},\bar{A}^1) \to H^2(B_{\boldsymbol{Q}(r)})$ need not be a monomorphism. However, $\mu_R{}^*x_{r+j}=0$ for $j=1,\cdots,n-r$ may be seen as follows. Consider for some j $(1\leqq j\leqq n-r)$ the $\boldsymbol{U}(1)$-bundle $\mathfrak{E}^{r+j}$ over $B_{\boldsymbol{Q}(n)}$; it admits over $\bar{A}$ the section θ_j. The map $\mu\colon B_{\boldsymbol{Q}(r)} \to B_{\boldsymbol{Q}(n)}$ induces over $B_{\boldsymbol{Q}(r)}$ a $\boldsymbol{U}(1)$-bundle ("counter-image of $\mathfrak{E}^{r+j}$") and explicit construction shows that the induced cross-section ψ over $\bar{A}^1$ (defined by $\psi(u)=(u,\theta_j\mu(u))$) can be extended all over $B_{\boldsymbol{Q}(r)}$. Thus, by naturality, $\mu_R{}^*x_{r+j}=0$ (from this $\mu_R{}^*S_R{}^{q+1}(x_1,\cdots,x_n)=0$ follows because of $q\geqq r$).

11.b. *Relative Pontryagin classes.*

From now on, the coefficients are integers mod p, where p is prime and >2. The following definition and naturality property would be valid without alteration with integer coefficients, but Whitney duality is not.

Let $\mathfrak{B}=(E,p,X,\boldsymbol{SO}(n))$ be a principal $\boldsymbol{SO}(n)$-bundle induced by a map $\tau\colon X\to B_{\boldsymbol{SO}(n)}$. Let $\sigma\colon B_{\boldsymbol{SO}(n)}\to B_{\boldsymbol{U}(n)}$ be the mapping corresponding to the inclusion $\boldsymbol{SO}(n)\to\boldsymbol{U}(n)$. Then $\sigma\circ\tau$ induces over X a principal $\boldsymbol{U}(n)$-bundle $\mathfrak{B}_C=(E_C,\pi,X,\boldsymbol{U}(n))$.

Let θ^r be a cross-section over the closed subset A of X in the bundle $\mathfrak{B}^r$ associated to $\mathfrak{B}$ with fibre $V_{n,n-2r+1}$. For $k\geqq r$, we have a cross-section θ^k over A in $\mathfrak{B}^k$ (with fibre $V_{n,n-2k+1}$) induced by θ^r. These cross-sections provide cross-sections $\bar{\theta}_C{}^k$ over A in the corresponding "complex" bundles $\mathfrak{B}_C{}^k$, associated to $\mathfrak{B}_C$ with fibre $W_{n,n-2k+1}$) as follows. Let $\bar{\tau}$, $\bar{\sigma}$ be the maps of the total space covering τ, σ respectively ($\bar{\sigma}$ is not a bundle map), and let h be the bundle map covering $\sigma\circ\tau$. Then h is injective (actually a homeomorphism) on each fibre $W_{n,n-2k+1}$, and $\theta_C{}^k$ may be defined by

$$(11.8)\qquad \pi\theta_C{}^k(a)=a,\qquad \theta_C{}^k(a)=h^{-1}\bar{\sigma}\bar{\tau}\theta^k(a)$$

Definition. The $4k$-dimensional relative Pontryagin class

$$P_R{}^k\in H^{4k}(X,X;\boldsymbol{Z}_p)\ \mathrm{mod}\ A$$

corresponding to the cross-section θ^r, defined for $k \geqq r$, is given by

(11.9) $$P_R^k = (-1)^k C_R^{2k},$$

where C_R^{2k} is the relative Chern class mod A of $\mathfrak{B}_C$ corresponding to θ_C^k.

Naturality. Let $\mathfrak{B}$, $\mathfrak{B}'$ be two $\boldsymbol{SO}(n)$-bundles over X and X' respectively, such that $\mathfrak{B}$ is induced from $\mathfrak{B}'$ by a map $f\colon X \to X'$. Let $\mathfrak{B}^k = (E^k, X)$, $\mathfrak{B}'^k = (E'^k, X')$ be the associated bundles with fibre $V_{n,n-2k+1}$ and assume that cross-sections θ^r, θ'^r in $\mathfrak{B}^r$, $\mathfrak{B}'^r$ are given over closed subsets $A \subset X$, $A' \subset X'$ respectively, such that $f(A) \subset A'$ and $\theta' f(a) = \bar{f}\theta(a)$, ($\bar{f}\colon E^k \to E'^k$ cover f) for every $a \in A$.

Let P^k and P'^k be the relative Pontryagin classes ($k \geqq r$) of $\mathfrak{B}$ and $\mathfrak{B}'$ corresponding to θ and θ' respectively. Then

(11.10) $$P^k = f^*(P'^k).$$

Proof. The above formula follows from the naturality property of relative Chern classes, provided we prove that $f_C\theta_C(a) = \theta'_C f(a)$, where $f_C\colon E_C{}^k \to E_C{}'^k$ covers f (Remark that $(E_C{}^k, X)$ is induced from $(E'_C{}^k, X')$ by $f\colon X \to X'$). We have

$$h\theta'_C f(a) = \overline{\sigma_7}\theta' f(a) = \overline{\sigma_7}\bar{f}\theta(a) = hf_C\theta_C(a)$$

and, since h is homeomorphic on the fibres,

$$\theta'_C f(a) = f\ \theta_C(a) \text{ for every } a \in A.$$

We are thus in position to apply the naturality property of relative Chern classes and (11.10) follows.

Whitney *duality* for Pontryagin classes will follow from Whitney duality for relative Chern classes reduced mod p. Because of the naturality property, we may restrict attention to the case of a classifying $\boldsymbol{SO}(n)$ bundle $\mathfrak{B} = (E_{\boldsymbol{SO}(n)}, p, B_{\boldsymbol{SO}(n)}, \boldsymbol{SO}(n))$ assuming a cross-section to be given over A (closed subset of $B_{\boldsymbol{SO}(n)}$) in the associated bundle $\mathfrak{B}^r$ with fibre $V_{n,n-2r+1}$. Let $\mathfrak{B}_C$ be the $\boldsymbol{U}(n)$-bundle over $B_{\boldsymbol{SO}(n)}$ induced by $\sigma\colon B_{\boldsymbol{SO}(n)} \to B_{\boldsymbol{U}(n)}$, and let $\theta_C{}^r$ be the cross-section over A induced by θ^r in the bundle $\mathfrak{B}_C{}^r$ associated to $\mathfrak{B}_C$ with fibre $W_{n,n-2r+1}$.

LEMMA (11.11). *The* $(4i+2)$*-dimensional relative Chern classes of* $\mathfrak{B}_C$ mod A *corresponding to* $\theta_C{}^r$ *are zero* mod p: $C_R{}^{2i+1} = 0$, $i \geqq r$.

Proof. Consider the diagram

$$(11.12)\qquad \begin{tikzcd} H^{4i+1}(E) \arrow[r, "\bar{i}^*"] & H^{4i+1}(\theta^i A) \arrow[r, "\delta"] \arrow[d, "\theta^*"] & H^{4i+2}(E^i, \theta^i A) \arrow[r] & H^{4i+2}(E^i) \\ H^{4i+1}(B_{SO(n)}) \arrow[u, "\omega^*"] \arrow[r, "i^*"] & H^{4i+1}(A) \arrow[r, "\delta"] & H^{4i+2}(B_{SO(n)}, A) \arrow[u, "\omega_R^*"] \arrow[r, "a^*"] & H^{4i+2}(B_{SO(n)}). \arrow[u, "\omega^*"] \end{tikzcd}$$

Since $a^*(C_R^{2i+1}) = C^{2i+1}$, we have (by [5], Proposition 25.4) $a^* C_R^{2i+1} = 0$ (coefficients mod p). Therefore $C_R^{2i+1} = \delta x$, $x \in H^{4i+1}(A; \mathbf{Z}_p)$, and since θ^* is an isomorphism, there exists a class $u \in H^{4i+1}(\theta^i A)$, with $\theta^* u = x$. We have

$$\delta u = \delta\theta^{*-1} x = \omega_R^* \delta x = \omega_R^* C_R^{2i+1} = 0.$$

The last equality follows from consideration of the diagram

$$\begin{tikzcd} H^{4i+2}(E^i, \theta^i A) & H^{4i+2}(E_C{}^i, \theta_C{}^i A) \arrow[l, "\bar{\sigma}^*"'] \\ H^{4i+2}(B_{SO(n)}, A) \arrow[u, "\omega_R^*"] & H^{4i+2}(B_{U(n)}, A). \arrow[l, "\sigma^*"'] \arrow[u, "\pi^*"'] \end{tikzcd}$$

Indeed, $\omega_R^* C_R^{2i+1} = \omega_R^* \sigma^* c_R^{2i+1} = \bar{\sigma}^* \pi^* c_R^{2i+1} = 0$ (by (11.4)).

By exactness of the rows in diagram (11.12), there exists an element $z \in H^{4i+1}(E^i)$ such that $\bar{i}^* z = u$. The assertion $C_R^{2i+1} = 0$ follows from the fact that ω^* is an *epimorphism* in every dimension (By [5], Theorem 23.2, $H^*(B_{SO(2m-1)}; \mathbf{Z}_p) = \mathbf{Z}_p[P_1, \cdots, P_{m-1}]$ and

$$H^*(B_{SO(2m)}; \mathbf{Z}_p) = \mathbf{Z}_p[P_1, \cdots, P_{m-1}, W_{2m}].$$

Notice that E^i is classifying space for $\mathbf{SO}(2i-1)$.).

Remark. Using the fact that ω^* is still an epimorphism if *integer* coefficients are used (see a forthcoming paper by A. Borel and F. Hirzebruch), the same method would give $2C_R^{2i+1} = 0$, where C_R^{2i+1} is the integer relative Chern class of a $\mathbf{U}(n)$-bundle obtained from an $\mathbf{SO}(n)$-bundle.

Whitney duality for relative Pontryagin classes is an immediate consequence of the same property for relative Chern classes with coefficients mod p, making use of Lemma (11.11).

Let $\mathfrak{B}_i$ be two $\mathbf{SO}(n_i)$-bundles $(i = 1, 2)$ over X and let θ_i be cross-sections over (closed subsets) $A_i \subset X$ in the associated bundles $\mathfrak{B}_i{}^{r_i} = (E_i{}^{r_i}, p_i, X, V_{n_i, n_i - 2r_i + 1})$.

Let $\mathfrak{B} = \mathfrak{B}_1 \oplus \mathfrak{B}_2$ be the Whitney sum and θ the cross-section over $A = A_1 \cap A_2$ in $\mathfrak{B}^r$ (with fibre $V_{n,n-2r+1}$, where $n = n_1 + n_2$, $r = r_1 + r_2$) obtained using $\theta_1^{r_1}$ and $\theta_2^{r_2}$.

P_R^k, the relative Pontryagin class of dimension $4k$ of corresponding to θ is defined for $k \geqq r$. Similarly, $P^1{}_R{}^k$, $P^2{}_R{}^k$ are defined for $k \geqq r_1$ and $k \geqq r_2$ respectively.

One has in $H^{4k}(X, A; \mathbf{Z}_p)$:

$$P_R^k = P^1{}_R{}^k + \cdots + P^1{}_R{}^{r_1} \cdot P^2{}_R{}^{k-r_1} + \cdots + P^2{}_R{}^k \qquad (k \geqq r). \tag{11.13}$$

Although some absolute classes might appear in the above formula, each cup-product contains at least one relative class. The sum on the right consists only of *relative* classes.

Pontryagin classes as symmetric functions. Consider again a principal $\boldsymbol{SO}(n)$-bundle $\mathfrak{B} = (E, p, X, \boldsymbol{SO}(n))$ and define the subgroup $\boldsymbol{Q}(n)$ of $\boldsymbol{SO}(n)$ by $\boldsymbol{Q}(n) = \boldsymbol{SO}(2) \times \boldsymbol{SO}(2) \times \cdots \times \boldsymbol{SO}(2)$ or $\boldsymbol{Q}(n) = \boldsymbol{SO}(2) \times \boldsymbol{SO}(2) \times \cdots \times \boldsymbol{SO}(2) \times \boldsymbol{SO}(1)$ according as $n = 2m$ or $n = 2m + 1$ (m factors $\boldsymbol{SO}(2)$ in both cases). Consider the quotient space $E/\boldsymbol{Q}(n)$. The principal fibre bundle $(E, \pi, E/\boldsymbol{Q}(n), \boldsymbol{Q}(n))$ is the Whitney sum of m principal $\boldsymbol{SO}(2)$-bundles $\mathfrak{E}^1, \mathfrak{E}^2, \cdots, \mathfrak{E}^m$. Let $x_1, x_2, \cdots, x_m$ be their Chern classes ($\boldsymbol{SO}(2)$ being identified with $\boldsymbol{U}(1)$).

Assuming a cross-section θ over $A \subset X$ (closed subset) to be given in $\mathfrak{B}^r = (E^r, p, X, V_{n,n-2r+1})$ we obtain cross-sections in $\mathfrak{E}^1$ for $i = r, r+1, \cdots, m$ over $\bar{A} \subset B_{\boldsymbol{Q}(n)}$ as follows: E^r is the base space of a principal $\boldsymbol{SO}(2r-1)$-bundle $(E, E^r, \boldsymbol{SO}(2r-1))$. Let $\bar{E}^r$ be the space of flags over $\bar{E}^r$, i.e. $\bar{E}^r = E/\boldsymbol{Q}(2r-1)$, where $\boldsymbol{Q}(2r-1)$ is the subgroup of $\boldsymbol{SO}(n)$ consisting of the matrices of the type

$$\begin{bmatrix} D_1 & & & & & & & \\ & \cdot & & & & & & \\ & & \cdot & & 0 & & & \\ & & & \cdot & & & & \\ & & & & D_{r-1} & & & \\ 0 & & & & 1 & & & \\ & & & & & 1 & & \\ & & & & & & \cdot & \\ & & & & & & & \cdot \\ & & & & & & & & \cdot \\ & & & & & & & & & 1 \end{bmatrix}, \text{ where } D_i = \begin{bmatrix} \cos(2\pi x_i) & \sin(2\pi x_i) \\ -\sin(2\pi x_i) & \cos(2\pi x_i) \end{bmatrix}.$$

We have the diagram

$$(11.14)\qquad \begin{array}{ccccc} & & \bar{E}^r & \xrightarrow{\omega} & \bar{X} \\ & & {\scriptstyle\rho_E}\downarrow & & \downarrow{\scriptstyle\rho_X} \\ A & \xrightarrow{\theta} & E^r & \xrightarrow{p} & X \end{array}$$

The map $\omega\colon \bar{E}^r \to \bar{X}$ is induced by the identity $E \to E(\boldsymbol{Q}(2r-1) \subset \boldsymbol{Q}(n))$. Let $\bar{A}$ be the space of the bundle $(\bar{A}, \rho_A, A, \boldsymbol{F}(2r-1), \boldsymbol{SO}(2r-1))$ induced by θ. Notations: $\boldsymbol{F}(2r-1) = \boldsymbol{SO}(2r-1)/\boldsymbol{Q}(2r-1)$, $\bar{\theta}\colon \bar{A} \to \bar{E}^r$ is the bundle map covering θ. The maps $p\theta$ and $\omega\bar{\theta}$ are injective and we may consider $\bar{A}$ as a subset of $\bar{X}$. Geometrically, a point of $\bar{A} \subset \bar{X}$ consists of a point a of A together with a sequence of m oriented 2-planes $\pi_1, \pi_2, \cdots, \pi_m$, such that π_r contains the first, π_{r+1} the second and the third, $\cdots$, π_m the $(n-2r)$-th and the $(n-2r+1)$-st vectors of $\theta(a)$, assuming $n = 2m$. If $n = 2m+1$, π_r contains the first and second, etc., π_m the $(n-2r)$-th and the $(n-2r+1)$-st vectors of $\theta(a)$. In both cases ($n = 2m$ or $2m+1$), since π_r is oriented, $\mathfrak{E}^r$ admits a cross-section over $\bar{A}$ and so do $\mathfrak{E}^{r+1}, \cdots, \mathfrak{E}^m$. We define $x_r, x_{r+1}, \cdots, x_m$ (the characteristic classes of $\mathfrak{E}^r, \cdots, \mathfrak{E}^m$) as *relative* classes corresponding to the cross-sections given by θ over $\bar{A}$. The elementary symmetric functions $S^k(x_1^2, \cdots, x_m^2)$ are then relative classes mod $\bar{A}$ (of dimension $4k$) for $k \geqq r$ and will, consequently, be denoted by $S_R^k(x_1^2, \cdots, x_m^2)$. One has the

THEOREM (11.15). *Let $\rho_R^*\colon H^*(X, A; \boldsymbol{Z}_p) \to H^*(\bar{X}, \bar{A}; \boldsymbol{Z}_p)$ be the homomorphism induced by $\rho_R\colon (\bar{X}, \bar{A}) \to (X, A)$. Then* (a) *$\rho_R^*$ is a monomorphism, and* (b) *$\rho_R^* P_R^k = S_R^k(x_1^2, \cdots, x_m^2)$ for $k \geqq r$.*

Proof of (a). Consider the diagram

$$\begin{array}{ccccccc} H^{4k-1}(\bar{X}) & \xrightarrow{\bar{i}^*} & H^{4k-1}(\bar{A}) & \xrightarrow{\delta} & H^{4k}(\bar{X}, \bar{A}) & \xrightarrow{\bar{a}^*} & H^{4k}(\bar{X}) \\ \uparrow{\scriptstyle\rho_X^*} & & \uparrow{\scriptstyle\rho_A^*} & & \uparrow{\scriptstyle\rho_R^*} & & \uparrow{\scriptstyle\rho_X^*} \\ H^{4k-1}(X) & \xrightarrow{i^*} & H^{4k-1}(A) & \xrightarrow{\delta} & H^{4k}(X, A) & \xrightarrow{a^*} & H^{4k}(X) \end{array}$$

(coefficients = remainders mod p, p = prime > 2), where ρ_X^* and ρ_A^* are monomorphisms in every dimension (see [5], Theorem 23.2) and i, $\bar{i}$ are the inclusions $i = p\theta$, $\bar{i} = \omega\bar{\theta}$.

The situation is entirely similar to the one in the proof of Theorem (5.1). A straightforward exactness argument shows that (a) follows from the Lemma: *If $b \in H^*(A)$ and $w \in H^*(\bar{X})$ are such that $\rho_A^* b = \bar{i}^* w$, then there exists a class $v \in H^*(X)$, such that $i^* v = b$.*

The proof of the lemma is entirely similar to the one given in the proof of Theorem (5.1).

Proof of (b). By naturality, it is sufficient to prove the formula $\rho_R{}^*P_R{}^k = S_R{}^k(x_1{}^2, \cdots, x_m{}^2)$ for the bundle $\mathfrak{B} = (B_{Q(n)}, \rho_n, B_{SO(n)}, \boldsymbol{F}(n), \boldsymbol{SO}(n))$ obtained from the classifying bundle ($B_{Q(n)}$ is the space of flags over $B_{SO(n)}$). The diagram (11.14) reads in this case.

$$\begin{array}{ccccc} A & \xrightarrow{\bar\theta} & B_{SO(2r-1)} & \xrightarrow{\mu} & B_{Q(n)} \\ \downarrow{\scriptstyle \rho_A} & & \downarrow{\scriptstyle \rho_{2r-1}} & & \downarrow{\scriptstyle \rho_n} \\ \bar A & \xrightarrow{\theta} & B_{Q(2r-1)} & \xrightarrow{p} & B_{SO(n)} \end{array}$$

and there is a similar diagram with k substituted for r for every k such that $r \leqq k \leqq \frac{1}{2}(n+1)$. Considering $\bar A$ as a subset of $B_{Q(2r-1)}$ and $B_{Q(n)}$ by the injections $\bar\theta$ and $\mu\bar\theta$, we have the following diagram

$$\begin{array}{ccccccc} H^{4k-1}(B_{Q(2k-1)}) & \xrightarrow{\bar\theta^*} & H^{4k-1}(\bar A) & \xrightarrow{\delta} & H^{4k}(_{Q(2k-1)}, \bar A) & \xrightarrow{a^*} & H^{4k}(B_{Q(2k-1)}) \\ \uparrow{\scriptstyle \mu^*} & & \updownarrow{\scriptstyle \cong} & & \uparrow{\scriptstyle \mu_R{}^*} & & \uparrow{\scriptstyle \mu^*} \\ H^{4k-1}(B_{Q(n)}) & \xrightarrow{\theta^*} & H^{4k-1}(\bar A) & \xrightarrow{\delta} & H^{4k}(B_{Q(n)}, \bar A) & \xrightarrow{a^*} & H^{4k}(B_{Q(n)}). \end{array}$$

From the relation $\rho^*(P^k) = S^k(x_1{}^2, \cdots, x_m{}^2)$ for the absolute Pontryagin classes, (see [6], Proposition 5.1), we have $a^*y = 0$, setting $y = \rho_R{}^*P_R{}^k - S_R{}^k(x_1{}^2, \cdots, x_m{}^2)$. Indeed, $a^*P_R{}^k = P^k$ is immediately seen from the definition and the corresponding equality for Chern classes.

By an exactness argument used several times in this paper, (b) follows from $\mu_R{}^*y = 0$. The proof of $\mu_R{}^*S_R{}^k(x_1{}^2, \cdots, x_m{}^2) = 0$ is similar to the one given for Chern classes. The equality $\mu_R{}^*\rho_R{}^*P_R{}^k = 0$ follows from consideration of the diagram

$$\begin{array}{ccc} H^{4k}(B_{Q(2k-1)}, \bar A) & \xleftarrow{\mu_R{}^*} & H^{4k}(B_{Q(n)}, \bar A) \\ \uparrow{\scriptstyle \kappa_R{}^*} & & \uparrow{\scriptstyle \rho_R{}^*} \\ H^{4k}(B_{SO(2k-1)}, A) & \xleftarrow{p^*} & H^{4k}(B_{SO(n)}, A). \end{array}$$

One has $\mu_R{}^*\rho_R{}^*P_R{}^k = \kappa_R{}^*p^*P_R{}^k$. Now, $p^*(P_R{}^k) = 0$ because

$$P_R{}^k = (-1)^k\sigma^*(C_R{}^{2k})$$

and of the diagram

$$\begin{array}{ccc} H^{4k}(B_{SO(2k-1)}, A) & \xleftarrow{\bar{\sigma}^*} & H^{4k}(B_{U(2k-1)}, \bar{\sigma}A) \\ \uparrow p^* & & \uparrow \pi^* \\ H^{4k}(B_{SO(n)}, A) & \xleftarrow{\sigma^*} & H^{4k}(B_{U(n)}, A) \end{array}$$

$p^*(P_R{}^k) = (-1)^k p^*\sigma^*(C_R{}^{2k}) = (-1)^k\bar{\sigma}^*\pi^*(C_R{}^{2k}) = 0$, by definition of the relative Chern classes.

This completes the proof of Theorem (11.15).

MASSACHUSETTS INSTITUTE OF TECHNOLOGY.

REFERENCES.

[1] J. Adem, "The iteration of the Steenrod squares in algebraic topology," *Proceedings of the National Academy of Science*, U. S. A., vol. 38 (1952), pp. 720-726.

[2] P. Alexandroff and H. Hopf, *Topologie*, Springer, 1935.

[3] A. Borel, "Sur la cohomologie des espaces fibrés principaux et des espaces homogènes de groupes de Lie compacts," *Annals of Mathematics*, vol. 57 (1953), pp. 115-207.

[4] ———, "La cohomologie mod 2 de certains espaces homogènes," *Commentarii Mathematici Helvetici*, vol. 27 (1953), pp. 165-197.

[5] ———, "Selected topics in the homology theory of fibre bundles," Mimeographed notes, Chicago, 1954.

[6] A. Borel and J. P. Serre, "Groupes de Lie et puissances réduites de Steenrod," *American Journal of Mathematics*, vol. 75 (1953), pp. 409-448.

[7] H. Cartan, "Sur l'itération des opérations de Steenrod," *Commentarii Mathematici Helvetici*, vol. 29 (1955), pp. 40-58.

[8] M. Kervaire, "Courbure intégrale généralisée et homotopie," *Mathematische Annalen*, vol. 131 (1956), pp. 219-252.

[9] J. Milnor, "On the immersion of n-manifolds in $(n+1)$-space," *Commentarii Mathematici Helvetici*, vol. 30 (1956), pp. 275-284.

[10] J. P. Serre, "Homologie singulière des espaces fibrés," *Annals of Mathematics*, vol. 54 (1951), pp. 425-505.

[11] N. Steenrod, "Cohomology invariants of mappings," *Annals of Mathematics*, vol. 50 (1949), pp. 954-988.

[12] ———, *Topology of fibre bundles*, Princeton Press, 1950.

[13] N. Steenrod and J. H. C. Whitehead, "Vector fields on the n-sphere," *Proceedings of the National Academy of Science*, vol. 37 (1951), pp. 58-63.

[14] R. Thom, "Quelques propriétés globales des variétés différentiables," *Commentarii Mathematici Hedvetici*, vol. 28 (1954), pp. 17-86.

[15] H. Whitney, "Differentiable manifolds," *Annals of Mathematics*, vol. 37 (1936), pp. 645-680.

[16] W. T. Wu, "Classes caractéristiques et i-carrés dans une variété," *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences*, Paris, vol. 230 (1950), pp. 508-511.

ON ARITHMETICAL SUMMATION PROCESSES.*

By Aurel Wintner.

Part I.

On the Eratosthenian summation process.

Let (C, p), (A), (L) and (E) refer to the summability of Σa_n (where $n = 1, 2, \cdots$) in the sense of Cesàro, Abel, Lambert and Eratosthenes respectively, the (E)-summability of Σa_n being defined by the existence of $\lim E_n$, where

$$E_n = E_n(a) = \sum_{k=1}^{n} \{n/k\} a_k \qquad \left(\text{or } E_n = \sum_{k=1}^{\infty} \{n/k\} a_k\right), \tag{1}$$

if $\{x\}$ an abbreviation for $[x]/x$ and $[x]$ denotes the greatest integer not exceeding x (so that the coefficient, $[n/k]k/n$, of a_k in (1) does not exceed 1, and is positive if $1 \leqq k \leqq n$). If $a_1, a_2, \cdots$ is replaced by $b_1 = b_1(a)$, $b_2 = b_2(a), \cdots$, where

$$b_n = \sum_{d \mid n} a_d/d, \tag{2}$$

then it is readily seen that, for every n,

$$M_n(b) = E_n(a), \tag{3}$$

where

$$M_n = M_n(b) = n^{-1} \sum_{k=1}^{n} b_k. \tag{4}$$

An arbitrary assignment of $a_1, a_2, \cdots$ is equivalent to an arbitrary assignment of $b_1, b_2, \cdots$, since the linear substitution (2) has a unique inverse; the explicit form of the latter is

$$a_n = \sum_{d \mid n} \mu(n/d) b_d/n, \tag{5}$$

where $\mu(k)$ is Möbius' factor. According to (3) and (4), a series Σa_n is (E)-summable if and only if

$$M(b) = \lim_{n \to \infty} (b_1 + \cdots + b_n)/n \tag{6}$$

* Received December 14, 1956.

exists, that is, if and only if the series

$$\sum_{n=1}^{\infty} \Delta b_n, \text{ where } \Delta b_n = b_n - b_{n-1}, \qquad (b_0 = 0), \tag{7}$$

is $(C, 1)$-summable.

The replacement of $a_1, a_2, \cdots$ by $E_1, E_2, \cdots$, occurring in particular cases, $a_n = \chi(n)/n$, in the *Disquisitiones Arithmeticae* (and, as a matter of fact, in Euler's writings; for further references, cf. [4], pp. 9-10), was propounded in [4] as a general summation method. First, the following facts were proved in [4], pp. 11-13:

(i) *The convergence of* Σa_n *is neither necessary nor sufficient for the* (E)*-summability of* Σa_n *(but the sum and the* (E)*-sum of* Σa_n *must have the same value when both exist).*

The first of the assertions of (i) can be formulated by saying that the (E)-method is not a regular summation method (in the sense of Toeplitz' criteria). In view of the first and second assertions of (i), there arose the question after appropriate Tauberian restrictions. In this regard, the following criteria were proved in [4], p. 21:

(ii) *The convergence of* Σa_n *is equivalent to the* (E)*-summability of* Σa_n *if any one of the following three conditions is satisfied:*

(I) $|na_n| < \text{const.}$; (II) $a_n \geqq 0$; (III) a_n is lacunary

(by (III) is meant that $a_n = 0$ unless n is in a sequence $m_1, m_2, \cdots$ satisfying $m_{k+1}/m_k > \text{const.}$ for some const. > 1).

The prime number theorem follows from that assertion of (ii) which assures that (I) and the (E)-summability of Σa_n together are sufficient for the convergence of Σa_n. In order to see this, it is sufficient to choose $a_n = \mu(n)/n$ and to observe that the (E)-summability of $\Sigma \mu(n)/n$ is trivial (in fact, the case $a_n = \mu(n)/n$ of (5) belongs to $b_1 = 1$, $b_2 = b_3 = \cdots = 0$; so that the numerator of (5) is independent of n, and so the limit (6) certainly exists).

In a subsequent paper, in which [4] is overlooked, Ingham [3] proved a result which refines those assertions of (ii) in which the convergence of Σa_n is the conclusion. In fact, any one of the three conditions (I), (II), (III) is sufficient in order that the $(C, 1)$-simmability, or just the (A)-summability, of Σa_n be equivalent to the convergence of Σa_n (Hardy, Littlewood, Hardy and Littlewood), whereas Ingham's relevant result is as follows:

(iii) *The* (E)*-summability of* Σa_n *is sufficient for the* $(C,1)$*-summability of* Σa_n *(to the same sum).*

Actually, Ingham also shows that the $(C,1)$ in (iii) can be improved to (C,ϵ) if $\epsilon > 0$ (in view of (i), this is not true for $\epsilon = 0$). Incidentally, Theorem (IX_1) in [4], p. 20, is nothing but Theorem 2 in [3], p. 175, since, according to Landau's extension of Hardy's $(C,1)$-lemma, the $(C,1)$-summability and the convergence of Σa_n are equivalent if $a_1 + \cdots + a_n > -\mathrm{const}$.

For reasons explained in [4], I referred to the replacement of the evaluation of Σa_n by means of the sequence $E_1, E_2, \cdots$ as an Eratosthenian summation method. But in the same way as [3], an Appendix to Hardy's posthumous book [1], pp. 376-380, overlooks [4] and renames the (E)-process Ingham's summation process. The oversight is the more curious as Hardy himself wrote the review of [4] in *Nature*, vol. 152 (1943), p. 708.

If

$$M(n) = \sum_{k=1}^{n} \mu(k), \tag{8}$$

then what is involved in the proof of (ii) in [4] as well as in the proof of (iii) in [3] or [1] is something like

$$M(n) = O(n/\log^2 n), \tag{9}$$

an estimate which assures the boundedness of the relevant "Lebesgue constants." This is more than the prime number theorem, which is just

$$M(n) = o(n). \tag{10}$$

But the distinction is less substantial than it was at that time, since, after Selberg, (9) is just as "elementary" as (10). In [3] or [1], the estimate (9) is used directly, and in [4] indirectly, via the theorem of Hardy and Littlewood according to which the (L)-summability of a series is always sufficient for its (A)-summability; cf. [1], pp. 373-374. In fact, something like (9), rather than just (10), is needed in the Hardy-Littlewood proof of $(L) \Rightarrow (A)$. The way in which [4] used $(L) \Rightarrow (A)$ in the proof of (ii) was the following:

The connection (2) between $a_1, a_2, \cdots$ and $b_1, b_2, \cdots$, a connection which, through (3), leads to the formulation (6) of the (E)-summability of Σa_n, is formally equivalent to

$$\sum_{n=1}^{\infty} b_n r^n = \sum_{n=1}^{\infty} n a_n (r^n + r^{2n} + \cdots). \tag{11}$$

8

It is also clear that the convergence of either of the series (11) for $0 \leqq r < r_0$ implies the convergence of the other series (to the same sum) for $0 < r < r_0$, provided that $0 < r_0 \leqq 1$ (this proviso is satisfied, with $r_0 = 1$, in all the cases considered above). But (11) can be written in the form

$$(12) \qquad \sum_{n=1}^{\infty} (\Delta b_n) r^n = (1 - r) \sum_{n=1}^{\infty} n a_n r^n / (1 - r^n),$$

where $\Delta b = b_n - b_{n-1}$ (with $\Delta b_1 = b_1$), and (12) shows that a series Σa_n is (L)-summable if and only if the corresponding series (7) is (A)-summable (to the same sum).

In what follows, (iii) will be verified, again with the aid of (9), along the lines followed in the proof of (i), that is, by showing that Toeplitz's conditions for a regular process are satisfied by the linear transformation, say

$$(13) \qquad S_n = \sum_{m=1}^{n} a_{nm} E_m,$$

where the coefficients a_{nm} are absolute constants and the sequences $E_1, E_2, \cdots$, $S_1, S_2, \cdots$, belonging to an arbitrary series Σa_n, are defined by (1) and

$$(14) \qquad (n+1) S_n = \sum_{m=1}^{n} \sum_{k=1}^{m} a_k$$

respectively. The assertion of (iii) is equivalent to

$$(15) \qquad \limsup_{n \to \infty} \sum_{m=1}^{n} | a_{nm} | < \infty.$$

In fact, the remaining two of three Toeplitz conditions are

$$(16) \qquad \lim_{n \to \infty} \sum_{m=1}^{n} a_{nm} = 1$$

and

$$(17) \qquad \lim_{m \to \infty} a_{nm} = 0 \text{ for } n = 1, 2, \cdots,$$

and both (16) and (17) are trivially satisfied in the present case (cf. below).

First, from (1) where $\{x\} = [x]/x$,

$$n E_n = \sum_{k=1}^{n} [n/k] k a_k; \text{ hence } n E_n = \sum_{m=1}^{n} \sum_{k=1}^{[n/m]} k a_k$$

or, if $T_n = \sum_{k=1}^{n} k a_k$,

$$n E_n = \sum_{m=1}^{n} T_{[n/m]} = \sum_{m=1}^{\infty} T_{[n/m]}$$

$(T_0 = 0)$, and so, by Möbius' inversion,

$$T_n = \sum_{j=1}^{\infty} \mu(j)\,[n/j]\,E_{[n/j]} = \sum_{j=1}^{n} \mu(j)\,[n/j]\,E_{[n/j]}.$$

But it follows from (4) that $S_n = \sum_{k=1}^{n} T_k/(k+k^2)$, since $T_k = \sum_{m=1}^{k} ma_m$. Consequently

$$S_n = \sum_{k=1}^{n} (k+k^2)^{-1} \sum_{j=1}^{k} \mu(j)\,[k/j]\,E_{[k/j]}.$$

If the last relation is thought of as written in the form (13), then it is seen that

$$(18) \qquad a_{nm} = m \sum_{k=1}^{n} (k+k^2)^{-1} \sum_{k/(m+1)<j\leq k/m} \mu(j).$$

Here the interior sum is $M([k/m]) - M([k/(m+1)])$, by (8). It follows therefore from $(k+k^2)^{-1} < k^{-2}$ that (15) is true if

$$\sum_{k=1}^{n} k^{-2} \sum_{m=1}^{n} m\,|\,M([k/m]) - M([k/(m+1)])\,| < \text{Const.}$$

holds for every n and for some constant. But the existence of such a constant is a straightforward consequence of (9) (if, according to Dirichlet's classical device, the sum is broken into two parts, those contributed by the summation ranges between 1 and $[n^{\frac{1}{2}}]$ and between $[n^{\frac{1}{2}}]+1$ and n).

The four summation processes, mentioned before (1), are "comparable" (in the sense of Hausdorff); their relative position is as follows:

$$(19) \qquad (\mathrm{E}) \Rightarrow (C,1) \Rightarrow (L) \Rightarrow (A).$$

In fact, $(\mathrm{E}) \Rightarrow (C,1)$ is equivalent to (15)-(17), that is, to Ingham's result (iii), and $(L) \Rightarrow (A)$ is the corresponding result of Hardy and Littlewood ([1], Theorem 260), finally $(C,1) \Rightarrow (L)$ is an elementary Abelian lemma of Hardy. The only thing that is peculiar about (E) is that, according to (i),

(20) both $(K) \Rightarrow (\mathrm{E})$ and $(\mathrm{E}) \Rightarrow (K)$ are false,

where (K) denotes convergence.

None of the arrows occurring in (19) can be inverted. First, since $(K) \Rightarrow (C,1)$, it is clear from (20) that $(C,1) \Rightarrow (\mathrm{E})$ is false. Next, in order to obtain a series Σa_n which disproves $(L) \Rightarrow (C,1)$, it is sufficient to choose $a_n = (-1)^n n$, since the corresponding function $\Sigma a_n z^n$, where $|z| < 1$, remains regular at $z = 1$ but $a_n \neq o(n)$. Finally, the irreversibility of the third arrow in (19) was proved by Hardy and Littlewood [2].

If $(\mathrm{E}) \Rightarrow (L)$ is concluded from (19), then there becomes involved the prime number theorem, since the latter is involved in $(\mathrm{E}) \Rightarrow (C, 1)$. Actually, the part $(\mathrm{E}) \Rightarrow (L)$ of (19) is nothing but Frobenius' $(C,1)$-variant of Abel's continuity theorem, to be applied to (7). In fact, (3), (4) and (12) show that the $(C,1)$-summability and the (A)-summability of (7) are equivalent, respectively, to the (E)-summability and the (L)-summability of Σa_n. The latter equivalences, when combined with that theorem of Hardy and Littlewood for which Karamata gave a simple proof, also yield the following Tauberian theorem (which does not involve the prime number theorem):

(†) *The* (E)*-summability of a (real) series* Σa_n *follows from its* (L)*-summability if* $\liminf_{n\to\infty} \sum_{d \mid n} a_d > -\infty$.

In fact, the numbers (2) are the partial sums of (7).

Whereas the corollary $(\mathrm{E}) \Rightarrow (L)$ of (19) is straightforward, the *weakest* corollary, $(\mathrm{E}) \Rightarrow (A)$, of the whole of (19), cannot be proved without the prime number theorem. For it is readily seen from (1) that if $a_n = \mu(n)/n$, then $E_n = 1/n$; hence $\lim E_n = 0$, and so the (E)-summability of Σa_n is now trivial. But if $(\mathrm{E}) \Rightarrow (A)$ is granted, then, since $|na_n| <$ const., the convergence of $\Sigma \mu(n)/n$ follows from Littlewood's (A)-Tauberian theorem.

It is clear from the proof of (15) that the first of the three implications (19) can be refined as follows:

(I) *There exists a universal constant* α *for which* $\operatorname{osc} S_n \leqq \alpha \operatorname{osc} E_n$ *holds whenever* $\limsup |E_n| < \infty$.

Here S_n and E_n are defined, in terms of an arbitrary Σa_n, by (14) and (1) respectively, and $\operatorname{osc} B_n$ denotes the upper limit of $|B_n - B_m|$ as $n > m \to \infty$.

The corresponding refinement of the third of the three implications (19) is equivalent to the following lemma:

(II) *There exists a universal constant* β *for which*

$$\operatorname{osc}(1-r)\sum_{n=1}^{\infty}\Big(\sum_{k=1}^{n} a_k\Big)r^n \leqq \beta \operatorname{osc}(1-r)\sum_{n=1}^{\infty}\Big(\sum_{k \mid n} a_k\Big)r^n \tag{21}$$

holds whenever $\limsup \big|(1-r)\sum_{n=1}^{\infty}\big(\sum_{k\mid n} a_k\big)r^n\big| < \infty$, *with* $\sum_{n=1}^{\infty} |a_n| r^n < \infty$ *for* $0 < r < 1$.

It is understood that $\operatorname{osc} f(r)$ denotes the upper limit of $|f(r) - f(s)|$

as $1 > r > s \rightarrow 1$. Hence an inspection of the Hardy-Littlewood proof of $(L) \Rightarrow (A)$ shows that, under the restrictions mentioned after (21),

$$\text{osc} \sum_{n=1}^{\infty} a_n r^n \leqq \beta \, \text{osc}(1-r) \sum_{n=1}^{\infty} n a_n r^n/(1-r^n) \tag{22}$$

holds for a certain universal β (the theorem of Hardy and Littlewood itself results if the osc on the right of (22) is assumed to be 0). But since the expression on the left of (21) is identical with the left of (22), it is seen from (2) and (11) that (21) is equivalent to the striking assertion of (II).

What is involved in (I) and (II) is (10), and more. I do not know whether (II) could be obtained as an "Abelian" consequence of (I) (somewhat along the lines of Tauber's own theorem, according to which (K) is equivalent to (A) and $T_n = o(n)$ together, where T_n is defined as in the proof of (15) above). Correspondingly, it would be worthwhile to determine the best values of the universal constants α, β occurring in (I), (II), and to decide whether $\alpha < \beta$.

There is a corresponding question concerning an absolute constant which belongs to $(\mathrm{E}) \Rightarrow (L)$ in the same way as α and β belong to $(\mathrm{E}) \Rightarrow (C,1)$ and $(L) \Rightarrow (A)$ respectively. Actually, the relevant analogue of (I) and (II) is straightforward. It is true that if $(\mathrm{E}) \Rightarrow (L)$ is concluded from (19), then (iii) becomes involved. But this is unnecessary. In fact, it turns out that what corresponds to (I) and (II) is

$$\text{osc}(1-r) \sum_{n=1}^{\infty} n a_n r^n/(1-r^n) \leqq \text{osc}\, E_n \tag{23}$$

(if $\limsup |E_n| < \infty$), and that the refinement (23) of $(\mathrm{E}) \Rightarrow (L)$ can be reduced to the O-version of Frobenius' Abelian implication $(C,1) \Rightarrow (A)$. This refinement of $(C,1) \Rightarrow (A)$, a refinement the truth of which is clear from the proof of $(C,1) \Rightarrow (A)$, states that

$$\text{osc} \sum_{n=1}^{\infty} a_n r^n \leqq \text{osc}\, S_n \tag{24}$$

(if $\limsup |S_n| < \infty$). But if the series Σa_n, to which (24) refers, is replaced by the series (7), defined by (2), then it is seen from (3), (4) and (12) that (24) can be written in the form (23).

Since $(C,1) \Rightarrow (\mathrm{E})$ and $(A) \Rightarrow (L)$ are false, what corresponds to (I) and (II) in the reverse direction is the order of infinity of the "Lebesgue constants." In [4], p. 12, the order of infinity was shown to be logarithmic for the transformation leading from (K) to (E).

Part II.

On Axerian theorems.

A related (but more elementary) arithmetical summation problem which is dealt with in Hardy's posthumous Appendix [1], referred to above, centers about Axer's lemma (cf., e. g., [5], p. 202). In what follows, both types of Axerian facts, presented in [1] (Theorem 267, p. 378) in a form communicated by Ingham, will be improved.

Let $s_1, s_2, \cdots$ be a sequence of numbers satisfying

$$\sum_{n \leqq x} s_n = o(x) \quad \text{and} \quad s_n = O(1), \tag{25}$$

and let $f(x)$, where $1 \leqq x < \infty$, be a function which is of bounded variation on every *finite* interval,

$$\int_1^x |df(u)| < \infty \quad \text{if} \quad 1 \leqq x < \infty. \tag{26}$$

By an Axerian condition (*) will be meant a restriction, to be imposed on f alone, to the following effect: Whenever conditions (26) and (*) are satisfied by f, the relation

$$\sum_{n \leqq x} f(x/n) s_n = o(x) \tag{27}$$

holds for every sequence $s_1, s_2, \cdots$ satisfying (25).

Such an Axerian condition (*) is contained in Landau's writings on the "equivalence" questions in the theory of primes (the most delicate of these questions is the deduction of

$$\sum_{n \leqq x} \mu(n)/n = o(1) \quad \text{from} \quad \psi(x) \sim x, \tag{28}$$

that is, from $\pi(x) \sim x/\log x$). What Landau actually proves is that the existence of a θ satisfying

$$f(x) = O(x^\theta), \quad \text{where} \quad \theta = \text{const.} < 1, \tag{29}$$

is a (*)-condition (in other words, that (27) follows from (25) whenever (26) and (29) are satisfied by f). In order to obtain (28), it is sufficient to apply (29), with $\theta = \frac{1}{2}$, to the $f(x)$ which represents Dirichlet's remainder term in the divisor problem. Cf. the references in [6], p. 121, and the reproduction of Landau's arguments in [1], pp. 378-380.

In what follows, it will be shown that the Landau's (*)-criterion (29) can be replaced by

(30) $$f(x) = O(\phi(x)),$$

where $\phi(x)$ is any non-negative function satisfying the following conditions:

(31) $$\phi(x) \text{ is monotone} \qquad (\phi \geqq 0)$$

and

(32) $$\int_1^\infty x^{-2}\phi(x)\,dx < \infty$$

(hence (31) means that $\phi(x)$ is non-increasing). In other words, the following Axerian lemma will be proved:

(*) *If $f(x)$ is any function satisfying* (26) *and* (30), *where $\phi(x)$ is some function subject to* (31) *and* (32), *then* (27) *is a consequence of* (25).

Remark. In view of (**)) below, it is worth mentioning that the conditions imposed on $f(x)$, $1 \leqq x < \infty$, by (*) can simply be formulated as follows: *If $g(x) = f(1/x)$, then*

$$\text{(I)} \quad \int_{+0}^1 \operatorname*{l.u.b.}_{x \leqq u \leqq 1} |g(u)|\,dx < \infty, \qquad \text{(II)} \quad \int_x^1 |dg(u)| < \infty,$$

where $0 < x \leqq 1$ (so that $\int_{+0}^1 |dg(x)| = \infty$ and $\limsup_{x \to +0} |g(x)| = \infty$ are allowed).

Needless to say, (31) and (32) are satisfied by $\phi(x) = x^\theta$ if $0 \leqq \theta < 1$, that is, when (30) becomes (29). The simplest case in which (29) fails to apply, but the italicized lemma is applicable, results if (30) holds for $\phi(x) = x^2/\log x$. In particular, it follows from (8) and (9) that the following fact is a corollary: *If $f(x) = \sum_{n \leqq x} \mu(n)$, then* (27) *holds for any sequence $s_1, s_2, \cdots$ satisfying* (25).*

---

* There is a corresponding deduction from Mertens' hypothesis, that is, from

(a) $$\sum_{n \leqq x} \mu(n) = O(\sqrt{x}),$$

as follows: The series

(b) $$\sum_{n=1}^\infty s_n \mu(n)/\sqrt{n}$$

is convergent for *every* sequence $s_1, s_2, \cdots$ satisfying

(c) $$s_1 \geqq s_2 \geqq \cdots \text{ and } \lim s_n = 0$$

if (and only if) Mertens' hypothesis is true.

In fact, (c) and the boundedness of the partial sums of a series Σr_n imply the

The nature of the restrictions imposed on $\phi(x)$ by (31) and (32) is of particular interest, since (31) and (32) are precisely the conditions underlying a theorem of Landau (Satz 3 in [6], pp. 123-124), a theorem which applies in a direction just the opposite to the direction of an Axerian implication (and *assumes,* therefore, the prime number theorem; in this regard, cf. Ingham's Theorems 1 and 3 in [3]).

The proof of the italicized criterion (*) will depend on the following elementary lemma: If $h(t)$ is any function which is monotone on the open half-line $0 < t < \infty$ and is integrable on its closure (in the sense that both

$$(31') \qquad \lim_{\epsilon \to 0} \int_{\epsilon}^{1} h(t)\,dt \text{ and } \lim_{\epsilon \to 0} \int_{1}^{1/\epsilon} h(t)\,dt$$

exist), then, as $x \to \infty$,

$$(32') \qquad \sum_{n=1}^{\infty} h(n/x) \sim Cx, \text{ where } C = \int_{+0}^{\infty} h(t)\,dt$$

(the convergence of the series (32′) for any fixed $x > 0$ is part of the statement). The fact that the monotony of $h(t)$ on $0 < t < \infty$ and the finiteness of both integrals (31′) imply (32′), is just a restatement of no. 30 in chap. II of [7].

Starting with a function $\phi(x)$ on the closed half-line $1 \leqq x < \infty$ and assuming (31) and (32), define on the open half-line $0 < t < \infty$ a function $h(t)$ by placing $h(t) = \phi(1/t)$ if $0 < t \leqq 1$ and $h(t) = 0$ if $1 < t < \infty$. Then it is clear that the conditions required of h in (32′), the convergence of both integrals (31′) and the monotony of h, are satisfied. Hence it is also clear that the same conditions remain satisfied if any positive $\epsilon < 1$ is chosen and if the preceding $h(t)$ is replaced by the function $h_\epsilon(t)$ which is

---

convergence of the series $\Sigma r_n s_n$ (Abel-Dedekind). Hence, if the latter series is identified with (b), the convergence of (b) follows from the assumptions (c) if

$$\text{(d)} \qquad \sum_{n \leqq x} \mu(n)/\sqrt{n} = O(1)$$

is true. But while a partial summation leads from (d) to (a), it is clear that (a) *in itself* fails to imply (d). It was however shown in [8] that, for *function-theoretical* Tauberian reasons, (d) must be true if (a) is true. Accordingly, (a) would be disproved if one could construct a *single* sequence $s_1, s_2, \cdots$ satisfying the conditions (c) and rendering the series (b) divergent.

Conversely, if there does not exist such a sequence, then (d), hence (a), is true. For, according to Hadamard's converse of the elementary lemma used above (Abel-Dedekind), the partial sums of a series Σr_n must be bounded if the series $\Sigma r_n s_n$ is convergent for *every* sequence $s_1, s_2, \cdots$ satisfying (c).

identical with the preceding $h(t)$ for $0 < t \leqq \epsilon$ and is 0 for $\epsilon < t < \infty$. Accordingly, (32′) is valid if $h(t)$ is replaced by $h_\epsilon(t)$. But then, since

$$\sum_{n=1}^{\infty} h_\epsilon(n/x) = \sum_{n \leqq \epsilon x} h_\epsilon(n/x) = \sum_{n \leqq \epsilon x} \phi(x/n),$$

(32′) reduces to $\sum_{n \leqq \epsilon x} \phi(x/n) \sim x \int_{+0}^{\infty} h_\epsilon(t)\,dt$, where ϵ is fixed $(0 < \epsilon < 1)$ and $x \to \infty$. But

$$\int_{+0}^{\infty} h_\epsilon(t)\,dt = \int_{+0}^{\epsilon} \phi(1/t)\,dt = \int_{1/\epsilon}^{\infty} u^{-2}\phi(u)\,du$$

(where $u = 1/t$), and the last integral tends to 0 as $\epsilon \to 0$, since (32) is assumed (and $\phi \geqq 0$). Consequently

$$\lim_{\epsilon \to 0} \lim_{x \to \infty} x^{-1} \sum_{n \leqq x} \phi(x/n) = 0, \tag{33}$$

where the existence of the interior lim (for fixed ϵ) and of the repeated limit are parts of the conclusion.

If the interior lim is replaced by lim sup, then it is clear from (30), where $\phi \geqq 0$, that ϕ can be replaced by $|f|$ in (33). It follows therefore from the second of the two assumptions (25) that

$$\lim_{\epsilon \to 0} \lim_{x \to \infty} x^{-1} \left| \sum_{n \leqq \epsilon x} f(x/n) s_n \right| = 0. \tag{34}$$

Note that neither (26) nor the first of the two assumptions (25) was used thus far.

The relation (34) turns out to be the main point in the proof of the assertion, (27), of (*). In fact, owing to (34), the balance of the proof of (27) becomes the same as it is under Landau's assumption (29). In fact, (35) takes the place of the third formula line in [1], p. 379, and the balance of the proof (p. 379, lines 1-10, in [1], where the delta is the present epsilon) can be repeated *verbatim*. Suffice it to say that the partial summation and its estimate, carried out in [1] on the second half of p. 379, involve nothing like (29) or (30) above, but merely (26) and (25). The proof of (*) is therefore complete.

Axer's own result (for references, cf. [6], p. 121) weakens the second of the two assumptions (25) so as to replace (25) by

$$\sum_{n \leqq x} s_n = o(x) \quad \text{and} \quad \sum_{n \leqq x} |s_n| = O(x), \tag{35}$$

and it states that (35) is sufficient for the truth of (27) if $f(x)$ is chosen as follows:

$$f(x) = x - [x] \text{ for } 1 \leqq x < \infty. \tag{36}$$

In Ingham's approach to the Axerian lemmas (Theorem 267 in [1]), the assumptions (25) and (35) are treated simultaneously (by a partial summation) and, correspondingly, it is proved that (27) can be concluded from (35) when (36) is generalized to any *bounded* function $f(x)$ satisfying (26).

It turns out however that, precisely under the assumption (35), the method of partial summations, which involves (26), is quite misleading. In fact, the true generalization of (36) (in the case (35)) has nothing to do with the assumption (26) of (*), since the true condition on $f(x)$, where $1 \leqq x < \infty$, is simply the following: If $g(x)$ is defined by

$$g(x) = f(1/x) \text{ if } 0 < x \leqq 1 \text{ and } g(x) = g(0) \text{ if } x = 0 \tag{37}$$

then

$$g(x) \text{ is } R\text{-integrable on } 0 \leqq x \leqq 1. \tag{38}$$

In contrast, what Ingham's result (quoted in the preceding section) requires is that $g(x)$ be bounded* on $0 < x \leqq 1$ and of bounded variation on $\epsilon \leqq x \leqq 1$ if $0 < \epsilon < 1$. Actually, that (38) alone is enough is clear from 4, p. 2, where a very short proof is given for Axer's own case, (36). In fact, a glance at that proof shows that only (38) is needed:

(**) *If $f(x)$, where $1 \leqq x < \infty$, is any function corresponding to which the condition* (38) *is satisfied by the function $g(x)$ defined for $0 \leqq x \leqq 1$ by* (37), *then* (27) *is a consequence of* (35).

It is seen from (37) that (27) can be written in the form

$$\sum_{n \leqq x} g(n/x) = o(x) \tag{39}$$

(the value $g(0)$, assigned in (37) in order to render (38) meaningful, is immaterial). On the other hand, if x is replaced by ax and bx in the first of the assumptions (37), it follows, by subtraction, that

$$\sum_{a < x \leqq b} s_n = o(x) \tag{40}$$

---

* This is *implied* by (38). Originally (in Riemann's own paper and, until comparatively recently, in all books) the boundedness was an explicit *additional* condition of R-integrability. Actually the redundancy of this proviso, usually attributed to Landau's *Einführung in die Differentialrechnung und Integralrechnung* (1934), was recognized already by Hölder, on p. 3 of his *Beiträge zur Potentialtheorie* (Tübingen dissertation, Stuttgart, 1882).

whenever $0 \leqq a < b \leqq 1$. But (40) means that (39) is true for the function $g(x)$ which is 1 or 0 according as x is or is not on the subinterval $a < x \leqq b$ of $0 \leqq x \leqq 1$. Thus, for reasons of distributivity, (39) is true whenever $g(x)$ is a step-function (with a finite number of jumps) on $0 \leqq x \leqq 1$. Hence it is seen from the second of the assumptions (35), which was not used thus far, that (39) is true whenever $g(x)$ has the following property: There belongs to every $\epsilon > 0$ a pair of step-functions $g_\epsilon(x)$, $g^\epsilon(x)$ in such a way that, on the one hand, $g_\epsilon \leqq g \leqq g^\epsilon$ for $0 \leqq x \leqq 1$ and, on the other hand, the integral of $g^\epsilon - g_\epsilon$ (over $0 \leqq x \leqq 1$) is less than ϵ. Since the existence of such a pair g_ϵ, g^ϵ (for every $\epsilon > 0$) is equivalent to (38), the proof of (**) is complete.

Appendix.

Let Σa_n, where $n = 1, 2, \cdots$, be a series satisfying $\limsup |a_n|^{1/n} \leqq 1$. This means that the power series $f(x) = \Sigma a_n e^{-nx}$, where $f(\infty) = a_0 = 0$, converges for $x > 0$, and $f(x) = O(e^{-x})$ as $x \to \infty$, where $f' = df/dx$. Let the additional hypotheses requiring the convergence (at $x = +0$) of the integrals

$$\int_0^\infty f'(x)\,dx, \quad \int_0^\infty |f'(x)|\,dx, \quad \int_0^\infty \phi(x)\,dx, \text{ where } \phi(x) = \underset{x \leqq y < \infty}{\text{l. u. b.}} |f'(y)|,$$

(and, in all three cases, $\int_0^\infty = \int_{+0}^\infty$) be referred to by A, $|A|$, $\|A\|$ respectively. Thus A means the (A)-summability (Abel), and $|A|$ the absolute (A)-summability (J. M. Whittaker), of Σa_n, and $\|A\|$ requires more than $|A|$. Suffice it to say that, while $|A|$ is compatible with $\limsup x\,|f'(x)| = \infty$, where $x \to 0$, it follows from $\|A\|$ that $xf'(x) \to 0$, since the *monotony* and the integrability (at $x = +0$) of $\phi(x)$ imply that $x\phi(x) \to 0$.

A theorem of Ananda-Rau [9] states that *the* $\|A\|$*-summability of* Σa_n *implies the* (L)*-summability of* Σa_n. For this theorem of Ananda-Rau, whose proof is quite involved, Hardy and Littlewood [2], p. 259, gave a shorter proof which, however, a note in Hardy's book [1], p. 376, recognizes to be fallacious. By claiming a generalization of Ananda Rau's theorem, I committed a corresponding blunder in [10], p. 686, assertion (β_0). Since the sufficiency of $\|A\|$-summability for L-summability is just stated, but not proved, in Hardy's book [1], p. 376 (Theorem 263), the only correct proof

available in the literature appears to be Ananda-Rau's complicated approach [9].

It turns out, however, that Ananda-Rau's theorem follows in an almost trivial fashion if recourse is had to the simple device used at the beginning of the proof of (*) above. In order to see this, the following lemma (§), which has nothing to do with questions of summability, will first be isolated.

(§) *Let $F(x)$, where $0 < x < \infty$, be a function which is R-integrable on every interval $\lambda \leqq x \leqq \mu$, where $0 < \lambda < \mu < \infty$, and suppose that there exists a monotone function $\phi(x)$ satisfying $|F(x)| \leqq \phi(x)$ and having a finite (improper) integral over $0 < x < \infty$. Then*

$$h \sum_{n=1}^{\infty} F(nh) \to \int_0^{\infty} F(x)dx \text{ as } h \to +0 \qquad \left(\int_0^{\infty} = \int_{+0}^{1} + \int_1^{\infty}\right)$$

(and $\sum_{n=1}^{\infty} F(nh)$ is absolutely convergent for every $h > 0$).

First, if F is monotone throughout (so that either F or $-F$ can be chosen to be ϕ), then the assertion of (§) becomes the straightforward fact referred to at the beginning of the proof of (*) above ([7], chap. II, no. 30). Next, if F satisfies only the assumptions made in (§), then, by the proof of (*) above,

$$\lim_{\lambda \to 0} \limsup_{h \to 0} |h \sum_{n \leqq \lambda h} F(nh)| = 0,$$

since $\phi(x)$ is a montone and integrable majorant of $|F(x)|$ near $x = +0$. But if $x = +0$ is replaced by $x = \infty$, then the same argument shows that

$$\lim_{\mu \to \infty} \limsup_{h \to 0} |h \sum_{n \geqq \mu h} F(nh)| = 0.$$

Finally, since $F(x)$ is supposed to be R-integrable on every interval $\lambda \leqq x \leqq \mu$, where $0 < \lambda < \mu < \infty$, it is clear that, when λ (> 0) and μ ($> \lambda$) are fixed,

$$h \sum_{\lambda h < n < \mu h} F(nh) \to \int_{\lambda}^{\mu} F(x)dx \text{ as } h \to 0.$$

Clearly, the assertion of (§) follows from the last three formula lines (by adding three epsilons).

In order to obtain Ananda-Rau's theorem as a corollary of (§), suppose that Σa_n is $\| A \|$-summable. By definition, this means that the assumptions

of (§) are satisfied by $F(x)=f'(x)$, where $f(x)=\Sigma a_n e^{-nx}$ and $f'=df/dx$. Hence, if h is replaced by x in the assertion of (§),

$$\lim_{x\to 0} x\sum_{n=1}^{\infty} f'(nx)=\int_{+0}^{\infty} f'(x)dx=f(\infty)-f(+0)=-f(+0)$$

(the last integral is [absolutely] convergent, since $\| A \|$-summability clearly implies [absolute (A)-summability, hence] (A)-summability, that is, the existence of $f(+0)$, whereas $f(\infty)=0$ is obvious from $f'(x)=O(e^{-x})$ as $x\to\infty$). On the other hand, it is clear that, if $0<x<\infty$,

$$\sum_{n=1}^{\infty} f'(nx)=-g(x), \text{ where } g(x)=\sum_{n=1}^{\infty} na_n e^{-nx}/(1-e^{-nx})$$

(simply because $f(x)=\Sigma a_n e^{-nx}$, hence $f'(x)=-\Sigma na_n e^{-nx}$). But the last two formula lines show that $-xg(x)\to -f(+0)$ as $x\to +0$, which, in view of the definition of $g(x)$, proves that Σa_n is (L)-summable (to the same sum, $f(+0)$, as under (A)-summability).

Clearly, the ϕ-assumption imposed on $F(x)$ in (§) plays the part of a Tauberian restriction. There are other such restrictions which are sufficient for the truth of the assertion of (§). For the sake of simplicity, such a criterion will now be formulated for $0<x\leqq 1$ (rather than, as in (§), for $0<x<\infty$), as follows:

(§§) *If $F(x)$, where, $0<x\leqq 1$, is a real-valued function which has a continuous derivative $F'(x)$ satisfying the Tauberian restriction*

$$\liminf_{x\to 0} x^2F'(x)>-\infty,$$

then

$$h\sum_{n=1}^{[1/h]} F(nh)\to\int_{+0}^{1} F(x)dx \text{ as } h\to 0$$

whenever the improper integral of $F(x)$ exists (it need not converge absolutely).

In fact, if the preceding reduction of Ananda-Rau's theorem to (§) is compared with the proof of that criterion of Hardy and Littlewood on (L)-summability which is Theorem 262 in [1], then it becomes clear that the proof of (§§) is contained between the lines of [1], pp. 375-376. In the particular case of Hardy and Littlewood, the $F(x)$ of (§§) becomes the first derivative of $f(x)=\Sigma a_n e^{-nx}$ (so that the Tauberian restriction of (§§) leads to f'').

Needless to say, the existence of continuous $F'(x)$ is used in the proof of (§§) only in the immediate vicinity of $x = +0$ (for $\epsilon \leqq x \leqq 1$, where $0 < \epsilon < 1$, the R-integrability of $F(x)$ is more than sufficient). In addition, the continuity of the derivative can readily be dispensed with.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

Part I.

[1] G. H. Hardy, *Divergent Series*, Oxford, 1949.

[2] G. H. Hardy and J. E. Littlewood, "Notes on the theory of series (XX): on Lambert series," *Proceedings of the London Mathematical Society*, vol. 41 (1936), pp. 257-270.

[3] A. E. Ingham, "Some Tauberian theorems connected with the prime number theorem," *Journal of the London Mathematical Society*, vol. 20 (1945), pp. 171-180.

[4] A. Wintner, *Eratosthenian Averages*, Baltimore, 1943.

Part II.

[5] E. Hecke, *Theorie der algebraischen Zahlen*, Leipzig, 1923.

[6] E. Landau, Über einige neuere Grenzwertsätze, *Rendiconti del Circolo Matematico di Palermo*, vol. 34 (1912), pp. 121-131.

[7] G. Pólya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis*, vol. 1, Berlin, 1925.

[8] A. Wintner, "A note on Mertens' hypothesis," *Revista de Ciencias*, (Lima) Ano L, 1949, pp. 181-184.

Appendix.

[9] K. Ananada-Rau, "On Lambert's series," *Proceedings of the London Mathematical Society*, vol. 19 (1921), pp. 1-20.

[10] A. Wintner, "The sum formula of Euler-Maclaurin and the inversions of Fourier and Möbius," *American Journal of Mathematics*, vol. 69 (1947), pp. 685-708.

POSITIVITÄTSBEREICHE IM R^n.*

Von Max Koecher.

1. **Einleitung.** In der vorliegenden Untersuchung werden gewisse Kegelbereiche des n-dimensionalen reellen Zahlenraumes—Positivitätsbereiche genannt—durch drei einfache Eigenschaften definiert und diese Bereiche dann systematisch untersucht. Aus den relativ schwachen Forderungen, die an diese Bereiche gestellt werden, können eine Reihe von teilweise überraschenden Eigenschaften abgeleitet werden, die bisher nur in Spezialfällen bekannt waren. Es soll dies hier an einem Beispiel erläutert werden.

Das einfachste nicht-triviale Beispiel von Positivitätsbereichen geben diejenigen symmetrischen Matrizen, die als Matrix einer positiv definiten quadratischen Form fester Variablenzahl aufgefasst werden können. Prozesse wie z. B. die Bildung der Determinante oder der inversen Matrix und die Konstruktion einer beim Übergang zu äquivalenten Matrizen invarianten Massbestimmung lassen sich auch in den allgemeinen Positivitätsbereichen durchführen. Damit ergeben sich Beispiele von Bereichen, die ähnlich einfache Eigenschaften wie die quadratischen Formen besitzen.

Es zeigt sich, dass für Positivitätsbereiche eine gemeinsame Zahlentheorie (in Analogie zu der der quadratischen Formen) entwickelt werden kann. Darüber soll jedoch an anderer Stelle berichtet werden.

In der Darstellung des Stoffes wurde so vorgegangen, dass zuerst im Abschnitt **2** Positivitätsbereiche in topologischen Vektorräumen mit Norm definiert werden und für diese eine Reihe von einfachsten Eigenschaften abgeleitet werden. Ab Abschnitt **3** tritt jedoch eine Beschränkung auf den n-dimensionalen Zahlenraum ein. An manchen Stellen mögliche Verallgemeinerungen wurden dabei ausser acht gelassen. Bei der Lektüre ist zu empfehlen, sich an den Beispielen in **11** zu orientieren.

2. **Positivitätsbereiche in topologischen Vektorräumen.** Es sei R der mit der natürlichen Topologie versehene Körper der reellen Zahlen und X ein Vektorraum über R. Wir setzen voraus, dass X eine Norm zulässt, d. h. es existiert eine Abbildung $x \to |x|$ von X in R^+ mit den folgenden Eigenschaften:

* Received March 9, 1957.

$$|x| = 0 \text{ ist äquivalent mit } x = 0,$$

$$|\lambda x| = |\lambda| \cdot |x| \text{ für alle } \lambda \text{ aus } R \text{ und } x \text{ aus } X,$$

$$|x + y| \leqq |x| + |y| \text{ für alle } x \text{ und } y \text{ aus } X.$$

Diese Norm definiert auf X eine Topologie und in dieser Topologie wird X zu einem topologischen Vektorraum über R.

Eine Abbildung $L(a,b)$ von $X \times X$ in R nennen wir zur Abkürzung eine *Bilinearform von* X, wenn gilt:

(B. 1) $L(a,b) = L(b,a)$ für alle a und b aus X,

(B. 2) $L(a,x)$ ist bei festem a in x linear und stetig.

Für eine solche Bilinearform L bezeichnen wir mit $X(L)$ die Menge der a aus X mit $L(a,x) = 0$ für alle x aus X. Wegen $L(0,x) = 0$, x aus X, ist $X(L)$ sicher nicht leer. Man sieht ausserdem sofort, dass $X(L)$ ein in X enthaltener topologischer Vektorraum über R ist. Eine Bilinearform L von X heisse *ausgeartet*, wenn $X(L)$ nicht nur aus der Null besteht.

Ist L eine Bilinearform von X, dann nennen wir eine Teilmenge Y von X einen *Positivitätsbereich von* X *mit* L *als Charakteristik*, wenn gilt:

(P. 0) Y ist offen und nicht leer.

(P. 1) Für alle a und b aus Y gilt $L(a,b) > 0$.

(P. 2) Liegt x nicht in Y, dann gibt es ein a aus $\bar{Y}$, welches nicht zu $X(L)$ gehört, mit $L(a,x) \leqq 0$.[1]

Man wird später an Beispielen sehen, dass L keineswegs durch die Punktmenge Y eindeutig bestimmt ist. Eine Punktmenge Y von X heist ein *Positivitätsbereich von* X schlechthin, wenn es eine Bilinearform L von X derart gibt, dass Y ein Positivitätsbereich von X mit L als Charakteristik wird. (P. 1) zeigt sofort, dass jede Charakteristik eines Positivitätsbereiches nicht identisch Null ist.

Offenbar sind die Axiome (P. 1) bzw. (P. 2) von den Axiomen (P. 0), (P. 2) bzw. (P. 0), (P. 1) unabhängig. Dass aber auch (P. 0) von (P. 1), (P. 2) unabhängig ist, zeigt das einfache Beispiel

$$X = R^2, Y = \{y : y = \begin{pmatrix} y^1 \\ y^2 \end{pmatrix}, y^1 \geqq 0, y^2 > 0\},$$

wenn man als Bilinearform das Skalarprodukt zweier Vektoren nimmt.

---

[1] Wie üblich wird für Teilmengen A von X mit $\bar{A}$ die abgeschlossene Hülle und mit $\mathring{A}$ der offene Kern von A bezeichnet und $Rd(A) = \bar{A} - \mathring{A}$ gesetzt.

Satz 1. *Für jeden Positivitätsbereich Y von X gilt:*

a) *Für jede Charakteristik L von Y und jedes a aus $\bar{Y}$, welches nicht in $X(L)$ liegt, ist $L(a,y) > 0$ für alle y aus Y.*

b) *Mit a und b liegt auch $a+b$ in Y.*

c) *Mit a liegt auch λa in Y, falls $\lambda > 0$.*[2]

d) *Liegt x nicht in $\bar{Y}$, dann gibt es a aus Y mit $L(a,x) < 0$.*

e) *Liegen a und $-a$ in $\bar{Y}$, dann gehört a für jede Charakteristik L zu $X(L)$.*

Beweis: a) Wegen (P. 1) kann man annehmen, dass a ein Punkt von $Rd(Y)$ ist. Man hat dann $L(a,y) \geqq 0$ für alle y aus Y. Nehmen wir jetzt an, dass $L(a,y) = 0$ für ein y aus Y erfüllt wäre. Da Y offen ist, gibt es zu jedem x aus X ein $\lambda > 0$, für das $y + \lambda x$ in Y liegt. Jetzt hat man

$$0 \leqq L(a, y+\lambda x) = L(a,y) + \lambda L(a,x) = \lambda L(a,x),$$

d.h. $L(a,x) \geqq 0$ für alle x aus X. Da mit x auch $-x$ zu X gehört, folgt $L(a,x) = 0$, d.h. a liegt in $X(L)$.

b) Würde für a und b aus Y der Punkt $a+b$ nicht in Y liegen, dann gibt es nach (P. 2) ein c aus $\bar{Y}$ mit

$$0 \geqq L(c, a+b) = L(c,a) + L(c,b)$$

und c nicht aus $X(L)$. Das ist aber ein Widerspruch, denn nach a) gilt $L(c,a) > 0$ und $L(c,b) > 0$.

c) Die Behauptung kann analog zu b) erschlossen werden.

d) Wir nehmen an, dass für ein x, welches nicht in $\bar{Y}$ liegt, die Behauptung falsch ist, d.h. es gilt $L(a,x) \geqq 0$ für alle a aus Y. Da das Komplement von $\bar{Y}$ sicher offen ist, können wir zu gegebenen c aus Y ein $\lambda > 0$ so bestimmen, dass auch $x + \lambda c$ nicht in $\bar{Y}$ liegt. Jetzt folgt für alle a aus Y nach (P. 1)

$$L(a, x+\lambda c) \geqq \lambda L(a,c) > 0$$

oder

$$L(a, x+\lambda c) \geqq 0 \text{ für alle } a \text{ aus } \bar{Y}.$$

Andererseits gibt es nach (P. 2) ein a aus $\bar{Y}$, welches nicht in $X(L)$ liegt, mit $L(a, x+\lambda c) \leqq 0$, d.h.

$$0 = L(a, x+\lambda c) = L(a,x) + \lambda L(a,c).$$

[2] Kleine griechische Buchstaben bezeichnen stets reelle Zahlen.

Da aber $L(a,x)$ und $L(a,c)$ nicht negativ sind, folgt $L(a,c)=0$ im Widerspruch zu a).

e) Nach (P.1) hat man sofort $L(a,y)\geqq 0$ und $L(-a,y)\geqq 0$ für alle y aus Y, d.h. $L(a,y)=0$. Zu jedem x aus X bestimmen wir nun wieder $\lambda>0$ derart, dass $y+\lambda x$ in Y liegt und erhalten

$$0=L(a,y+\lambda x)=L(a,y)+\lambda L(a,x)=\lambda L(a,x).$$

x ist also Punkt von $X(L)$ und der Satz ist vollständig bewiesen.

Als einfache Folge dieses Satzes ergibt sich, dass $X(L)$ für jede Charakteristik L von Y eine Teilmenge von $Rd(Y)$ ist. Denn ein Punkt a von $X(L)$ gehört wegen (P.1) nicht zu Y und wegen Satz 1d sicher zu $\bar{Y}$.

Die Positivitätsbereiche von topologischen Vektorräumen sind maximal, man überlegt sich nämlich sofort, dass für zwei Positivitätsbereiche Y und $\tilde{Y}$ von X mit gleicher Charakteristik aus $Y\subseteq\tilde{Y}$ stets $Y=\tilde{Y}$ folgt.

Ein Positivitätsbereich Y möge *ausgeartet* genannt werden, wenn Y eine ausgeartete Charakteristik besitzt.

SATZ 2. *Sei Y ein Positivitätsbereich von X.*

a) *Ist X direkte Summe zweier Vektorräume $\tilde{X}$ und X_0, $X=\tilde{X}\oplus X_0$, und gilt gleichzeitig eine Zerlegung $Y=\tilde{Y}\oplus X_0$, dann ist jede Charakteristik von Y ausgeartet.*

b) *Ist Y ausgeartet, dann gibt es eine Zerlegung $X=\tilde{X}\oplus X_0$ von X in zwei topologische Vektorräume, hierbei ist ausserdem $Y=\tilde{Y}\oplus X_0$ und $\tilde{Y}$ ist Positivitätsbereich von $\tilde{X}$.*

Beweis: a) Sei L eine Charakteristik von Y und $x=\tilde{x}+x_0$ die direkte Zerlegung des allgemeinen Punktes von X. Sind a und y jetzt zwei Punkte von Y in der gleichen Zerlegung, dann folgt nach (P.1)

$$0<L(a,y)=L(\tilde{a},\tilde{y})+L(\tilde{a},y_0)+L(a_0,\tilde{y})+L(a_0,y_0).$$

Für $a_0\to 0$ und $\tilde{y}\to 0$ hat man

$$L(\tilde{a},y_0)\geqq 0 \text{ für alle } \tilde{a} \text{ aus } \tilde{Y} \text{ und } y_0 \text{ aus } X_0,$$

und $\tilde{a}\to 0$, $\tilde{y}\to 0$ liefert entsprechend

$$L(a_0,y_0)\geqq 0 \text{ für alle } a_0 \text{ und } y_0 \text{ aus } X_0.$$

Zusammen ist also $L(a,y_0)\geqq 0$, d.h.

$$L(a,y_0)=0 \text{ für alle } a \text{ aus } Y \text{ und } y_0 \text{ aus } X_0.$$

Wie im Beweis zu Satz 1e folgt hieraus, dass dies sogar für alle a aus X richtig ist, d. h. X_0 ist Teilmenge von $X(L)$ und L ist sicher ausgeartet.

b) Wir zeigen zuerst, dass X in eine direkte Summe zerlegt werden kann. Wir wählen dazu ein c mit $|c| = 1$ aus $X(L)$, setzen $X_0 = \{\lambda c\}$ und definieren für x_0 aus X_0

$$M(x_0) = \lambda, \text{ falls } x_0 = \lambda c.$$

$M(x)$ ist dann eine auf X_0 definierte stetige Linearform, die nach dem Satz von Hahn-Banach auf ganz X als stetige Linearform fortgesetzt werden kann. Für x aus X setzt man

$$\tilde{x} = x - M(x) \cdot c, \qquad x_0 = M(x)c$$

Die Zerlegung $x = \tilde{x} \oplus x_0$ ist dann direkt und es gilt $M(\tilde{x}) = 0$. Man erhält also eine Zerlegung von X in eine direkte Summe zweier Vektorräume $\tilde{X}$ und X_0. Da X_0 von c aufgespannt wird, ist stets

$$L(a, x) = L(a, \tilde{x}) = L(\tilde{a}, \tilde{x}).$$

Wir setzen jetzt $\tilde{Y} = Y \cap \tilde{X}$ und zeigen, dass $\tilde{Y}$ Positivitätsbereich von $\tilde{X}$ mit $L(a, b)$ als Charakteristik ist. Zum Beweis dieser Behauptung ist zuerst (P. 0) trivial. Für (P. 1) seien $\tilde{a}$ und $\tilde{b}$ Punkte aus $\tilde{Y}$, dann gehören sie auch zu Y und man hat $L(\tilde{a}, \tilde{b}) > 0$. Ist andererseits $\tilde{x}$ nicht aus $\tilde{Y}$, dann liegt $\tilde{x}$ auch nicht in Y. Nach (P. 2) gibt es a aus $\bar{Y}$, welches nicht in $X(L)$ liegt, mit $0 \geqq L(a, \tilde{x}) = L(\tilde{a}, \tilde{x})$ und $\tilde{a}$ ist natürlich in $\bar{\tilde{Y}}$ enthalten. Damit ist aber auch (P. 2) für $\tilde{Y}$ nachgewiesen. Zum vollständigen Beweis der Behauptung fehlt noch der Nachweis von $Y = \tilde{Y} \oplus X_0$. Wie vorher überlegt man sich sofort, dass $\tilde{Y} \oplus X_0$ ein Positivitätsbereich mit L als Charakteristik ist und $Y \supseteq \tilde{Y} \oplus X_0$ gilt. Wegen der Maximalität der Positivitätsbereiche folgt daraus die Gleichheit.

Nach diesem Satz kann man die Positivitätsbereiche in einfacher Weise auf nicht ausgeartete Positivitätsbereiche zurückführen und man übersieht alle Positivitätsbereiche, wenn man die nicht ausgearteten kennt.

Ist Y ein nicht ausgearteter Positivitätsbereich von X, dann kann man für die Punkte von X eine Relation "$\geqq$" (bzw. "$>$") durch

$$x \geqq y \text{ ist äquivalent mit } x - y \in \bar{Y}$$

$$(\text{bzw. } x > y \text{ ist äquivalent mit } x - y \in Y)$$

definieren. Diese Relation induziert dann in X eine *Halbordnung*, die mit R verträglich ist, d. h. es gilt:

$$\text{Aus } x \geqq y \text{ und } y \geqq z \text{ folgt } x \geqq z.$$

$$\text{Aus } x \geqq y \text{ und } y \geqq x \text{ folgt } x = y.$$

$$\text{Aus } x \geqq y \text{ folgt } a + x \geqq a + y.$$

$$\text{Aus } x \geqq y \text{ und } 0 < \lambda \in R \text{ folgt } \lambda x \geqq \lambda y.$$

Wir werden in *Zukunft* bei nicht ausgearteten Positivitätsbereichen für deren Punkte y *auch* $y > 0$ schreiben. Für a und b aus X und y aus Y hat man $L(a,y) > L(b,y)$ wenn $a > b$, d.h. die Bilinearform ist monoton.

3. Positivitätsbereiche im R^n. Von nun an werden wir nur endlichdimensionale Vektorräume zu betrachten haben und können uns daher auf den Fall $X = R^n$ beschränken. Die Punkte x des R^n schreiben wir als *Spalten*

$$x = \begin{pmatrix} x^1 \\ x^2 \\ \cdot \\ \cdot \\ \cdot \\ x^n \end{pmatrix}, x' = (x^1, x^2, \cdots, x^n)$$

und benützen konsequent den Matrizenkalkül. Als Norm des R^n nehmen wir $|x| = (x'x)^{\frac{1}{2}}$. Eine Bilinearform $L(a,b)$ des R^n hat dann die Form $L(a,b) = a'Sb = b'Sa$ mit einer geeigneten reellen symmetrischen n-reihigen Matrix S. Wir wollen jetzt auch S eine Charakteristik von Y nennen, wenn dies für $L(a,b)$ der Fall ist. $L(a,b)$ ist genau dann ausgeartet, wenn $|S| = 0$. Aus Satz 2 folgt, dass such jeder ausgeartete Positivitätsbereich Y in der Form $Y = \bar{Y} \oplus R^q$ $(0 < q < n)$ schreiben lässt, wobei $\bar{Y}$ ein geeigneter Positivitätsbereich des R^{n-q} ist. Man erkennt hieraus, dass durch Y der Rang der Charakteristik S eindeutig bestimmt ist.

Für reelle Matrix W von n *Zeilen* und Spalten und Punktmenge A des R^n möge mit WA die Menge der Bildpunkte von A bei der Abbildung $y \to Wy$ bezeichnet werden. Es ist dann sofort zu sehen, dass mit Y auch WY, $|W| \neq 0$, ein Positivitätsbereich ist. Ist hier S eine Charakteristik von Y, dann ist $W'^{-1}SW^{-1}$ eine Charakteristik von WY.

Ist Y ein Positivitätsbereich des R^n, S eine Charakteristik von Y, dann kann man als Ergänzung zu Satz 1 zeigen, dass zu jedem a aus $Rd(Y)$ ein b aus $Rd(Y)$ mit $Sb \neq 0$ und $a'Sb = 0$ existiert. Sei dazu a_ν eine gegen a konvergente Folge, deren Glieder nicht aus $\bar{Y}$ gewählt sind. Nach Satz 1d gibt es dann b_ν aus Y mit $a'_\nu Sb_\nu < 0$. Man kann hier noch ohne Ein-

schränkung $|Sb_\nu|=1$ annehmen. Wir wählen nun eine Teilfolge der a_ν derart, dass die zugehörigen b_ν gegen einen Punkt b von $\bar{Y}$ konvergieren. Man hat dann $a'Sb \leqq 0$, $|Sb|=1$. Nach (P. 1) ist aber nur $a'Sb=0$ möglich.

4. Nicht ausgeartete Positivitätsbereiche des R^n. Da man sich nach den Überlegungen in **2** und **3** bei den weiteren Untersuchungen auf nicht ausgeartete Positivitätsbereicche beschränken kann, wollen wir von nun an stets annehmen, dass Y ein nicht ausgearteter Positivitätsbereich des R^n ist. S bezeichne immer eine Charakteristik von Y, nach Satz 2 ist dann $|S|$ von Null verschieden.

Lemma 1. *Zu jedem Kompaktum K von Y gibt es ein $\rho(K)=\rho(K;S)>0$ mit*

$$a'Sy \geqq \rho(K)\,|y|$$

für alle a aus K und alle y aus $\bar{Y}$.

Beweis: Mann kann ohne Einschränkung $y \neq 0$ voraussetzen und wegen der Homogenität beider Seiten der Behauptung genügt es, wenn das Lemma für alle a aus K und y aus $\bar{Y}$ mit $|y|=1$ bewiesen wird. Da dies eine kompakte Punktmenge in $\bar{Y}\times\bar{Y}$ darstellt, nimmt $a'Sy$ als stetige Funktion ihr Minimum a'_0Sy_0 in einem Punkte (a_0, y_0) an. Nach Satz 1a ist dieses Minimum aber positiv.

Wendet man Lemma 1 auf $K=\{a\}$, a aus Y, an, dann erhält man

$$a'Sy \geqq \rho(a)\,|y| \text{ für alle } y \text{ aus } \bar{Y}$$

und wir wollen—um die Bezeichnung eindeutig zu fixieren—stets

$$\rho(a)=\inf\{a'Sy : y \text{ aus } \bar{Y}, |y|=1\}$$

setzen.

In **2** hatten wir zu jedem Positivitätsbereich eine Halbordnung definiert. Wir können jetzt zeigen, dass im vorliegenden Falle diese Halbordnung archimedisch ist.

Lemma 2. *Zu jedem $a>0$ und x aus R^n gibt es $\lambda>0$ mit $\lambda a>x$.*

Beweis: Es genügt, wenn die Existenz eines λ mit $\lambda a \geqq x$ gezeigt wird, denn dann ist $(\lambda+1)a>x$. Wäre diese Behauptung aber falsch, d. h. gibt es a aus Y und x aus R^n mit $\lambda a - x \notin \bar{Y}$ für alle $\lambda>0$, dann gibt es zu jedem λ ein c_λ mit

$$c'_\lambda S(\lambda a - x)<0, \quad |c_\lambda|=1, \quad c_\lambda \text{ aus } Y.$$

Jetzt ist mit Lemma 1

$$0 < \rho(a) \leqq c'_\lambda Sa/\lambda < c'_\lambda Sx \leqq |\, Sx \,|/\lambda$$

und das ist für $\lambda \to \infty$ ein Widerspruch.

Man kann im Anschluss an Lemma 2 leicht zeigen, dass sich jedes x in der Form $x = a - b$ mit a, b aus Y schreiben lässt. Man hat zu gegebenen $a > 0$ und x nur λ so zu bestimmen, dass $b = \lambda a - x$ in Y liegt. Die nicht ausgearteten Positivitätsbereiche des R^n sind also auch Positivitätsbereiche im Sinne von N. Bourbaki.

Lemma 1 hat aber noch weitere Konsequenzen. *Z.B.* ist sofort zu sehen, dass die y aus $\bar{Y}$ mit $a \geqq y$ $(a > 0)$ eine beschränkte Punktmenge bilden. Ist weiter y_ν eine Folge aus $\bar{Y}$, für die bei geeigneten $a > 0$ die Folge $a'Sy_\nu$ konvergiert, dann ist y_ν von selbst beschränkt. Für beliebiges a ist dies natürlich keineswegs richtig.

Geometrisch kann man $\rho(a)$ so interpretieren, dass $\rho(a)$ ein Mass für den Abstand des Punktes a von $Rd(Y)$ angibt. Um dies hinreichend allgemein formulieren zu können, sagen wir, dass eine Folge a_ν uneigentlich nach Unendlich konvergiert $(a_\nu \to \infty)$, wenn $|\, a_\nu \,|$ über alle Grenzen wächst, und nennen ausserdem a_ν (uneigentlich) gegen einen Randpunkt von Y konvergent $(a_\nu \to Rd(Y))$, wenn $a_\nu \to \infty$ und $|\, a_\nu \,|^{-1} a_\nu$ gegen einen Randpunkt konvergiert. Jetzt kann man zeigen, dass eine (eigentlich oder uneigentlich) konvergente Folge a_ν aus Y mit $\rho(a_\nu) \to 0$ stets gegen einen Randpunkt konvergiert. Ist nämlich $a_\nu \to a \neq \infty$ und $\rho(a_\nu) \to 0$ und wäre a kein Randpunkt, dann ist $\rho(a) > 0$. Andererseits gibt es aber eine Folge y_ν aus $\bar{Y}$ mit $\rho(a_\nu) = a'_\nu S y_\nu$, $|\, y_\nu \,| = 1$, und man kann $y_\nu \to y$ annehmen. Es folgt $\rho(a_\nu) \to a'Sy \geqq \rho(a)$ und das ist ein Widerspruch. Im Falle $a_\nu \to \infty$ ist $b_\nu = |\, a_\nu \,|^{-1} a_\nu$ eigentlich konvergent. Wegen

$$\rho(b_\nu) = |\, a_\nu \,|^{-1} \rho(a_\nu) \to 0$$

folgt die Behauptung aber aus dem bereits bewiesenem Teil. Für eigentlich konvergente Folgen kann man diese Aussage umkehren.

5. Automorphismengruppe und Norm eines Positivitätsbereiches. Für nicht ausgeartete Positivitätsbereiche Y bezeichnen wir mit $\Sigma(Y)$ die multiplikative Gruppe der umkehrbaren reellen n-reihigen Matrizen W mit $W\bar{Y} = \bar{Y}$ und nennen $\Sigma(Y)$ die (lineare) *Automorphismengruppe von* Y. Sie ist offenbar unabhängig von der Auswahl der Charakteristik S von Y.

Für W aus $\Sigma(Y)$ und Charakteristik S von Y setzen wir $W^* = S^{-1}W'S$ und erhalten

LEMMA 3. *Mit W gehört auch W^* zu $\Sigma(Y)$.*

Beweis: Wir zeigen zuerst $W^*\bar{Y} \subseteq \bar{Y}$ für alle W aus $\Sigma(Y)$. Wäre das nicht richtig, dann gibt es ein y aus $\bar{Y}$, für das W^*y nicht in $\bar{Y}$ liegt. Nach Satz 1d gibt es dann a aus Y mit $0 > a'SW^*y = (Wa)'Sy$. Da aber Wa in Y liegt, ist das ein Widerspruch zu (P.1). Ersetzt man in der bereits bewiesenen Behauptung aber W durch W^{-1}, dann folgt schon $W^*\bar{Y} = \bar{Y}$, d. h. W^* liegt in $\Sigma(Y)$.

Eine auf Y definierte Funktion $f(y)$ möge nun eine *Norm von Y* heissen, wenn die folgenden Axiome erfüllt sind:

(N.1) $f(y)$ ist auf Y stetig,

(N.2) $f(y) > 0$ für alle y aus Y,

(N.3) $f(Wy) = \| W \| f(y)$ für alle W aus $\Sigma(Y)$.

Dabei bedeutet $\| W \|$ den absoluten Betrag der Determinante $| W |$ von W. Die Existenz wenigstens einer Norm für jedes Y wird sich später zeigen. Da für $\lambda > 0$ und Einheitsmatrix E die Matrix λE zu $\Sigma(Y)$ gehört, zeigt (N.3), dass jede Norm eine homogene Funktion von Grade n ist. Ebenfalls wegen (N.3) ergibt sich trivial, dass

$$dv = dy/f(y), \qquad dy = \prod_{k=1}^{n} dy^k$$

ein bei den Abbildungen $y \to Wy$, W aus $\Sigma(Y)$, invariantes Volumenelement darstellt. Es ist leicht zu sehen, dass umgekehrt Y sicher dann eine Norm besitzt, wenn es ein in diesem Sinne invariantes Volumenelement besitzt.

Wir wollen nun zeigen, dass jedes Y wenigstens eine Norm zulässt und definieren dazu für y aus Y

$$N(y) = \omega/M(y), \qquad M(y) = \int_Y e^{-y'St}\, dt$$

mit einer positiven Konstanten ω, die wir erst später festlegen wollen. Da es zu jedem Kompaktum K von Y nach Lemma 1 ein $\rho(K) > 0$ mit $y'St \geqq \rho(K) | t |$ für alle y aus K und t aus $\bar{Y}$ gibt, ist die Konvergenz des Integrals $M(y)$ sichergestellt und $N(y)$ ist in Y sogar *reell-analytisch.* (N.1) und (N.2) für $N(y)$ sind dann trivial. Ist nun W aus $\Sigma(Y)$ gegeben, dann substituiert man in $M(Wy)$

$$t = W^{*-1}x, \qquad dt = \| W \|^{-1}\, dx.$$

Wegen Lemma 3 erhält man als Integrationsgebiet wieder Y und hat daher auch (N.3) nachgewiesen.

Die Norm $N(y)$ ist also reell-analytisch und ausserdem noch monoton, denn wir hatten schon gesehen, dass aus $a > b > 0$ immer $a'St > b'St$ folgt und das liefert $N(a) > N(b)$.

Zum Schluss wollen wir noch einige Eigenschaften von $\Sigma(Y)$ zusammenstellen. Es wird sich dabei zeigen, dass diejenige Untergruppen von $\Sigma(Y)$, deren Elemente W der Bedingung $\| W \| = 1$ genügen, abgeschlossen ist.

Lemma 4. *Es sei W_ν eine Folge von Matrizen aus $\Sigma(Y)$.*

a) *Ist $W_\nu \to W$ konvergent, dann liegt W für $|W| \neq 0$ in $\Sigma(Y)$, im Falle $|W| = 0$ gilt $W\bar{Y} \subseteq Rd(Y)$ und es ist $W = 0$ falls $Wa = 0$ für ein a aus Y.*

b) *Ist $W_\nu a$ für ein a aus Y beschränkt, dann ist auch W_ν beschränkt.*

Beweis: a) Man hat nach Definition $W_\nu \bar{Y} = \bar{Y}$ und $W_\nu^{-1}\bar{Y} = \bar{Y}$. Da aber wegen $|W| \neq 0$ auch die Folge W_ν^{-1} konvergiert, folgt $W\bar{Y} \subseteq \bar{Y}$ und $W^{-1}\bar{Y} \subseteq \bar{Y}$.

Im Falle $|W| = 0$ gilt sicher $W\bar{Y} \subseteq \bar{Y}$. Wäre jetzt $W\bar{Y}$ keine Teilmenge von $Rd(Y)$, dann gibt es a aus Y mit $Wa \in Y$. Wegen (N. 1, 3) folgt dann

$$0 < N(Wa) = \lim_{\nu \to \infty} N(W_\nu a) = \lim_{\nu \to \infty} \| W_\nu \| N(a) = 0,$$

also ein Widerspruch.

Für die letzte Aussage von a) zeigen wir zuerst, dass auch $W^*a = 0$. Wäre nämlich $W^*a \neq 0$, dann ist nach Satz 1a

$$0 < a'SW^*a = (Wa)'Sa = 0.$$

Ist nun y beliebig aus Y gewählt, dann liegt Wy in $\bar{Y}$. Wäre $Wy \neq 0$, dann wieder nach Satz 1a

$$0 < a'SWy = (W^*a)'Sy = 0.$$

Es ist also $Wy = 0$ für alle y aus Y, d. h. $W = 0$.

b) Für alle y aus Y mit $|Sy| \leqq 1$ folgt nach Lemma 1

$$y'SW_\nu a > \rho(a) \, | W^*_\nu y |.$$

Andererseits sind die Elemente von $W_\nu a$ beschränkt, d. h. $|W_\nu a| \leqq \gamma_1$ und man hat[3]

$$| W^*_\nu y | \leqq | Sy | \, | W_\nu a | / \rho(a) \leqq \gamma_2$$

[3] γ_1, γ_2 usw. bezeichnen positive von ν unabhängige Zahlen.

oder $y'W^{*\prime}_\nu W^*_\nu y \leqq \gamma_3$. Jetzt ergibt sich aber leicht, dass dann die Elemente von W^*_ν und daher auch von W_ν beschränkt sind.

In ähnlicher Weise kann man noch zeigen, dass für jedes a aus Y die Gruppe der W aus $\Sigma(Y)$ mit $Wa = a$ kompakt ist.

6. Ein invariantes Linienelement von Y. Es sei von nun an stets $N(y)$ die spezielle reell-analytische Norm von Y nach **5**. Ein positiv definites Linienelement

$$ds^2 = \sum_{k,l} g_{kl}(y)\, dy^k dy^l$$

möge *bei $\Sigma(Y)$ invariant* heissen, wenn die durch ds definierte Massbestimmung von Y bei den Bewegungen $y \to Wy$, W aus $\Sigma(Y)$, invariant ist. Bekanntlich ist dies dann und nur dann der Fall, wenn

$$W' \cdot G(Wy) \cdot W = G(y), \qquad G(y) = (g_{kl}(y))$$

für jedes W aus $\Sigma(Y)$ erfüllt ist.

Satz 3. *Das Linienelement*

$$ds^2 = \sum_{k,l} h_{kl}(y)\, dy^k dy^l, \qquad h_{kl}(y) = -\partial^2 \log N(y)/\partial y^k \partial y^l$$

ist positiv definit und bei $\Sigma(Y)$ invariant.

Beweis: Die Invarianz dieses Linienelementes folgt trivial aus (N.3). Es muss lediglich gezeigt werden, dass $H(y) = (h_{kl}(y))$ in jedem Punkt y von Y als Matrix einer quadratischen Form positiv definit ist. Setzt man dazu $z = St$ und

$$M_k = \partial M(y)/\partial y^k = -\int_Y e^{-y'z} z^k\, dt,$$

$$M_{kl} = \partial^2 M(y)/\partial y^k dy^l = \int_Y e^{-y'z} z^k z^l\, dt,$$

dann folgt für x aus R^n

$$\begin{aligned}
2[M(y)]^2 x'H(y)x &= 2\sum_{k,l} x^k x^l (M \cdot M_{kl} - M_k \cdot M_l) \\
&= \sum_{k,l} x^k x^l (M \cdot M_{kl} + M \cdot M_{lk} - 2M_k M_l) \\
&= \sum_{k,l} x^k x^l \int_Y dt_1 \int_Y dt_2\; e^{-y'(z_1+z_2)} (z_1{}^k z_1{}^l + z_2{}^k z_2{}^l - 2z_1{}^k z_2{}^l) \\
&= \int_Y dt_1 \int_Y dt_2 e^{-y'(z_1+z_2)} [(x'z_1)^2 + (x'z_2)^2 - 2(x'z_1)(x'z_2)] \\
&= \int_Y dt_1 \int_Y dt_2\, e^{-y'(z_1+z_2)} [x'S(t_1 - t_2)]^2
\end{aligned}$$

und daher ist $x'H(y)x > 0$ für $x \neq 0$, d. h. $H(y)$ positiv definit.

Aus einem invarianten Linienelement kann man bekanntlich eine *invariante Metrik* konstruieren. Sei dazu C eine ganz in Y liegende Kurve mit der Parameterdarstellung $y=y(\tau)$, $\alpha \leqq \tau \leqq \beta$, bei der die Komponenten von $y(\tau)$ stückweise stetig differenzierbare Funktionen der (reellen) Variablen τ sind. Wie üblich definiert man dann die *Länge von* C in der durch das Linienelement nach Satz 3 gegebenen Massbestimmung durch

$$\lambda(C) = \left| \int_\alpha^\beta (\dot{y}'H(y)\dot{y})^{\frac{1}{2}} d\tau \right|.$$

Da ds bei $\Sigma(Y)$ invariant ist, folgt sofort $\lambda(WC) = \lambda(C)$ für jedes W aus $\Sigma(Y)$. Setzt man ausserdem für a und b aus Y

$$\lambda(a,b) = \inf\{\lambda(C) : y(\alpha) = a, y(\beta) = b, C \subset Y\}.$$

dann überlegt man sich sofort, dass $\lambda(a,b)$ eine bei $\Sigma(Y)$ invariante Metrik von Y darstellt, d. h. den folgenden Forderungen genügt:

(M. 1) $\lambda(a,b) = \lambda(b,a) \geqq 0$, $\lambda(a,b) = 0$ nur bei $a = b$,

(M. 2) $\lambda(a,b) \leqq \lambda(a,c) + \lambda(c,b)$ für a, b und c aus Y,

(M. 3) Es gibt zu jedem Kompaktum K von Y zwei Konstanten $\gamma > 0$ und $\gamma' > 0$ mit $\gamma \mid a-b \mid \leqq \lambda(a,b) \leqq \gamma' \mid a-b \mid$ für alle a und b aus K, d. h. die durch $\lambda(a,b)$ in Y induzierte Topologie ist mit der natürlichen Topologie des R^n äquivalent,

(M. 4) Für jedes W aus $\Sigma(Y)$ ist $\lambda(Wa, Wb) = \lambda(a,b)$.

Zwei Probleme drängen sich in diesem Zusammenhang auf:

(A) Ist es richtig, dass für Positivitätsbereiche die geodätischen Linien der hier angegebenen Massbestimmung eindeutig bestimmt sind?

(B) Ist die Metrik $\lambda(a,b)$ *vollständig*, d. h. strebt für jede Folge a_ν auf Y, die gegen einen Randpunkt von Y konvergiert, der Abstand $\lambda(a, a_\nu)$, a aus Y, gegen Unendlich?

Eine Klärung dieser beiden Fragen wäre von grossem Interesse.

7. Positivitätsbereiche mit Involution. Wie bisher sei S eine Charakteristik von Y und $W^* = S^{-1}W'S$. Wir hatten schon gesehen, dass mit W auch W^* in $\Sigma(Y)$ liegt.

Eine Abbildung $y \rightarrow y^*$ von Y auf sich soll eine *Involution von* Y genannt werden, wenn für alle y aus Y gilt:

(I.1) $(y^*)^* = y$,

(I.2) $N(y)N(y^*) = \text{const.}$,

(I.3) $(Wy)^* = W^{*-1}y^*$ für alle W aus $\Sigma(Y)$.

In diesem Abschnitt wollen wir zeigen, dass jeder nicht ausgearteter Positivitätsbereich wenigstens eine Involution zulässt.

Für eine auf Y definierte und dort einmal partiell differenzierbare Funktion $f(y)$ setzt man wie üblich

$$\operatorname{grad}_y f(y) = \begin{bmatrix} \partial f(y)/\partial y^1 \\ \cdot \\ \cdot \\ \cdot \\ \partial f(y)/\partial y^n \end{bmatrix}.$$

Zu jedem y aus Y definieren wir jetzt ein y^* durch

$$y^* = S^{-1} \operatorname{grad}_y \log N(y).$$

Man sieht sofort, dass sich die Euler'sche Differentialgleichung für die homogene Funktion $N(y)$ in der Form $y'Sy^* = n$ schreiben lassen. Sei weiter für alle y aus Y

$$H(y) = -(\partial^2 \log N(y)/\partial y^k \partial y^l), \qquad K(y) = -(\partial y^{*k}/\partial y^l).$$

dann erkennt man sofort die Richtigkeit von $H(y) = SK(y)$ und daher ist stets $|K(y)| \neq 0$, denn $H(y)$ war nach Satz 3 positiv definit. Man kann also auch y als Funktion der Komponenten von y^* auffassen und erhält dann leicht $K(y)K(y^*) = E$. Man beachte hier aber, dass bis jetzt keineswegs bewiesen ist, dass y^* auch zu Y gehört. Durch Differentiation von $y'Sy^* = n$ erhält man schliesslich noch $y^* = K(y)y$ für alle y aus Y.

SATZ 4. *Die Abbildung $y \to y^*$ ist eine Involution von Y. Man kann $N(y)$ so normieren, dass es ein c aus Y mit $N(c)N(c^*) = 1$ gibt.*

Nach dieser Normierung hat man an Stelle von (I.2) $N(y)N(y^*) = 1$ für alle y aus Y.

Beweis: Wir teilen den Beweis in mehrere Schritte auf:

1) *Beweis von* (I.3): Sei W aus $\Sigma(Y)$, $Wy = z$, dann folgt nach Definition

$$\begin{aligned}(Wy)^* &= S^{-1} \operatorname{grad}_z \log N(z) = S^{-1} \operatorname{grad}_z \log N(y)\\ &= S^{-1}W'^{-1} \operatorname{grad}_y \log N(y) = S^{-1}W'^{-1}Sy^*\end{aligned}$$

und das ist die Behauptung.

2) *Zu jedem a aus Y gibt es ein b aus Y mit* $a=b^*$. Man fragt dazu für gegebenes a aus Y nach dem Minimum der Linearform $a'Sy$ unter den Nebenbedingungen $N(y)=\eta$ $(\eta>0)$ und y aus $\bar{Y}$. Dieses Minimum wird in einem Punkt b von $\bar{Y}$ angenommen und nach der Methode von Lagrange erhält man als Bestimmungsgleichungen für b

$$a=\lambda b^*, \qquad N(b)=\eta.$$

Nach der ersten Gleichung liegt also entweder b^* oder $-b^*$ in Y. Wegen $b'Sb^*=n$ kann aber nur b^* ein Punkt von Y sein. Wegen 1) kann man jetzt aber $a=(1/\lambda\, b)^*$ schreiben und das ist die Behauptung

3) *Für alle a und y aus Y ist* $a'Sy^*\geqq 0$. $N(y)$ war nach **5** reell-analytisch, d. h. $\log N(y+\lambda a)$ ist in eine gewöhnliche Potenzreiche nach Potenzen von λ entwickelbar:

$$\log N(y+\lambda a)=\log N(y)+\lambda a'Sy^*+\cdots,$$

d. h. $\log\{N(y+\lambda a)/N(y)\}=\lambda a'Sy^*+\cdots$. Nach **5** war $N(y)$ ausserdem monoton, d. h. für $\lambda>0$ gilt $N(y+\lambda a)>N(y)$. Für $\lambda\to 0$ folgt dann aber schon die behauptete Ungleichung.

4) $y\to y^*$ *ist Abbildung von Y auf sich*. Nach 2) erhält man als Bild von Y bei dieser Abbildung eine Punktmenge, die Y umfasst. Da diese Abbildung stetig ist, genügt es offenbar, wenn wir noch zeigen, dass die Bildmenge in $\bar{Y}$ enthalten ist. Würde jetzt aber für ein y aus Y das zugehörige y^* nicht in $\bar{Y}$ liegen, dann gibt es nach Satz 1d ein a aus Y mit $a'Sy^*<0$. Nach 3) ist dies aber ein Widerspruch.

5) *Beweis von* (I. 1, 2): Nach den bereits bewiesenen Rechenregeln hat man

$$(y^*)^*=K(y^*)y^*=K(y^*)K(y)y=y.$$

Zum Beweis von (I. 2) setzen wir $f(y)=N(y)N(y^*)$ und erhalten

$$\begin{aligned}\operatorname{grad}_y \log f(y) &= \operatorname{grad}_y \log N(y) - K'(y)\operatorname{grad}_{y^*}\log N(y^*)\\ &= Sy^* - K'(y)S(y^*)^* = SK(y)(y-y^{**})=0,\end{aligned}$$

denn $SK(y)$ war symmetrisch.

6) *Normierung von* $N(y)$: Man gibt sich ein a aus Y vor und setzt in der Bezeichnung von **5**

$$M(a)=\alpha^n, \qquad M(a^*)=\beta^n, \quad \alpha>0, \quad \beta>0.$$

Bei $c = \sqrt{\alpha/\beta}\,a$ folgt $M(c) = M(c^*)$. Die Konstate ω in der Definition von $N(y)$ fixieren wir jetzt durch $\omega = M(c) = M(c^*)$ und erhalten die noch fehlende Behauptung.

Die hier definierten Matrizen $K(y)$ sind über die bisherigen Zusammenhange hinaus deswegen von besonderem Interesse, weil in allen bekannten Beispielen die folgende Frage positiv beantwortet werden kann:

(C) Gehört für jedes y aus Y die Matrix $K(y)$ zu $\Sigma(Y)$?

Eine allgemeine Lösung dieses Problems würde für Positivitätsbereiche einen Einblick in die Automorphismengruppe ermöglichen. In Richtung auf (C) kann man jedoch vorläufig nur beweisen, dass die folgenden drei Aussagen gleichbedeutend sind:

(1) Für jedes y aus Y gehört $K(y)$ zu $\Sigma(Y)$.

(2) Für alle a und y aus Y und hinreichend kleines $\lambda > 0$ gilt
$$y^* \geqq (y + \lambda a)^*.$$

(3) Für alle a und b aus Y mit $a > b$ gilt stets $b^* > a^*$.

Der Beweis dieser Äquivalenz soll übergangen werden.

Ausser den bisher erwähnten Eigenschaften der Involution $y \rightarrow y^*$ ist noch bemerkenswert, dass diese Involution mit der Metrik $\lambda(a,b)$ verträglich ist, d. h. es gilt $\lambda(a^*, b^*) = \lambda(a,b)$ für je zwei Punkte a und b von Y. Diese Invarianz ist leicht nachzuweisen, wenn man beachtet, dass für eine Parameterdarstellung $y(\tau)$ einer Kurve C von Y das Bild C^* von C bei der Abbildung $y \rightarrow y^*$ die Parameterdarstellung $y^*(\tau)$ hat. Ohne Schwierigkeiten zeigt man damit $\lambda(C) = \lambda(C^*)$ und der Übergang zum Infimum liefert die obige Behauptung.

8. Homogene Positivitätsbereiche. Einen nicht ausgearteten Positivitätsbereich Y nennen wir *homogen*, wenn es zu je zwei Punkten a und b aus Y ein W aus $\Sigma(Y)$ mit $a = Wb$ gibt. Es soll von jetzt immer vorausgesetzt werden, dass Y ein homogener Positivitätsbereich ist. Wegen (N. 3) in **5** ist eine Norm von Y dann bis auf einen konstanten Faktor eindeutig bestimmt.

Das Verhalten von $N(y)$ auf dem Rand von Y kann man vollständig übersehen, denn es gilt:

LEMMA 5. *Ist $a_\nu \rightarrow a$, a aus $Rd(Y)$, eine (eigentlich) konvergente Folge aus Y, dann konvergiert $N(a_\nu)$ gegen Null.*

Beweis: Wir wählen einen Punkt c von Y und bestimmen W_ν aus $\Sigma(Y)$ mit $a_\nu = W_\nu c$. Da die a_ν beschränkt sind, sind nach Lemma 4b die Matrizen W_ν beschränkt und wir können ohne Einschränkung annehmen, dass W_ν gegen eine Matrix W konvergiert. Wäre hier $|W| \neq 0$, dann liegt W nach Lemma 4a in $\Sigma(Y)$ und a könnte kein Randpunkt sein. Es ist also $|W| = 0$ und wegen $N(a_\nu) = \| W_\nu \| N(c)$ folgt die Behauptung.

Nach diesem Lemma ist also $N(y)$ stetig auf $\bar{Y}$ fortsetzbar und wir bezeichnen diese Fortsetzung wieder mit $N(y)$.

Man könnte in Analogie zu diesem Lemma vermuten, dass für jede gegen Unendlich konvergierende Folge auch die Norm nach Unendlich konvergiert. Dies ist jedoch nicht allgemein richtig, es gilt vielmehr nur

LEMMA 6. *Für jede Folge $a_\nu \geqq a > 0$, die gegen Unendlich konvergiert, strebt auch $N(a_\nu)$ gegen Unendlich.*

Beweis: Mit geeigneten W_ν aus $\Sigma(Y)$ kann man $a_\nu = W_\nu a$ schreiben, d. h. es ist $a \geqq W_\nu^{-1} a$, und die Vektoren $W_\nu^{-1} a$ sind beschränkt. Wegen Lemma 4b sind dann wieder die Matrizen W_ν^{-1} beschränkt und man kann ohne Einschränkung $W_\nu^{-1} \rightarrow W$ annehmen. Wäre hier aber $|W| \neq 0$, dann würde auch die Folge W_ν konvergieren. Da dies der Voraussetzung widerspricht, ist $|W| = 0$ und das ist die Behauptung.

Wir kommen nun zu zwei wichtigen Eindeutigkeitssätzen. Es zeigt sich nämlich, dass ein Positivitätsbereich, wenn er homogen ist, durch seine Norm eindeutig bestimmt ist.

SATZ 5. *Sind Y_1 und Y_2 zwei homogene Positivitätsbereiche mit den Normen $N_1(y)$ und $N_2(y)$ und gilt*

$$Y_1 \cap Y_2 \neq \emptyset,\ N_1(y) = N_2(y) \text{ für alle } y \text{ aus } Y_1 \cap Y_2,$$

dann ist $Y_1 = Y_2$.

Beweis: Es genügt, wenn wir

$$Rd(Y_1) \cap Y_2 = Rd(Y_2) \cap Y_1 = \emptyset$$

nachweisen, denn da Y_1 und Y_2 offene konvexe Mengen sind, folgt hieraus der Satz. Zum Beweis dieser Behauptung setzen wir für y aus $Y_1 \cup Y_2$

$$N(y) = N_1(y), \text{ falls } y \text{ aus } Y_1, \quad N(y) = N_2(y), \text{ falls } y \text{ aus } Y_2.$$

Nach Voraussetzung ist diese Definition eindeutig. Würde es nun einen Punkt a aus $Rd(Y_1) \cap Y_2$ geben, dann wählt man eine Folge a_ν aus $Y_1 \cap Y_2$, die gegen a konvergiert. Jetzt folgt aber

$$0 < N_2(a) = \lim_{\nu \rightarrow \infty} N_2(a_\nu) = \lim_{\nu \rightarrow \infty} N_1(a_\nu) = 0$$

unter Verwendung von Lemma 5. Dieser Widerspruch löst sich nur, wenn die obige Behauptung richtig ist.

Nach diesem Eindeutigkeitssatz ist es jetzt möglich, die Automorphismengruppe $\Sigma(Y)$ auf eine neue Art zu charakterisieren. Es zeigt sich nämlich, dass $\Sigma(Y)$ genau aus den "Einheiten" der Funktion $N(y)$ besteht.

SATZ 6. *Eine reelle nicht ausgeartete Matrix W gehört dann und nur dann zu $\Sigma(Y)$, wenn $WY \cap Y \neq \emptyset$ und $\| W \| \cdot N(W^{-1}y) = N(y)$ für alle y aus $WY \cap Y$.*

Beweis: Dass die W aus $\Sigma(Y)$ den angeführten Bedingugen genügen, ist sicher richtig. Sei also umgekehrt W eine reelle umkehrbare Matrix mit den angegebenen Eigenschaften. Wegen der Homogenität von $N(y)$ können wir ohne Einschänkung $\| W \| = 1$ annehmen. Bei $\tilde{Y} = WY$ hat man nach Voraussetzung

$$Y \cap \tilde{Y} \neq 0, \qquad N(W^{-1}y) = N(y) \text{ für alle } y \text{ aus } Y \cap \tilde{Y}.$$

$\tilde{Y}$ ist wieder homogener Positivitätsbereich und wir können als Norm $\tilde{N}(y) = N(W^{-1}y)$, y aus $\tilde{Y}$, nehmen. Nach Voraussetzung ist dann aber $N(y) = \tilde{N}(y)$ für alle y aus $Y \cap \tilde{Y}$ und Satz 5 ergibt $Y = \tilde{Y}$, d.h. W ist ein Element von $\Sigma(Y)$.

In **5** hatte man gesehen, dass das Volumenelement $dv = (dy)/N(y)$ bei den Abbildungen $y \to Wy$, W aus $\Sigma(Y)$, invariant ist. Da man andererseits ein Volumenelement der durch $H(y)$ (vergl. **6**) definierten Massbestimmung in der Form $\sqrt{|H(y)|}\,dy$ erhält und dies für homogene Positivitätsbereiche bis auf einen konstanten Faktor eindeutig bestimmt ist, folgt $N^2(y) \cdot |H(y)| = \alpha$. Nach den Überlegungen von **7** hat man jetzt

$$dv^* = \frac{1}{N(y^*)}\, dy^* = \| K(y) \| N(y) dy = \frac{\alpha}{\| S \| \; N(y)}\, dy = \frac{\alpha}{\| S \|}\, dv$$

Erstetzt man in dieser Gleichung y durch y^*, so folgt $\alpha = \| S \|$ und man hat

$$dv^* = dv, \qquad |H(y)| = \| S \| / N^2(y)$$

nachgewiesen.

9. Die Fläche $N(y) = 1$. Ist Y wieder ein homogener Positivitätsbereich, dann wollen wir unter der *Normfläche von Y* die Menge der y aus Y mit $N(y) = 1$ verstehen. Wegen der Homogenität von $N(y)$ ist diese Normfläche zusammenhängend.

Wir werden zeigen, dass sich die Normfläche für $y \to \infty$ überall an den

Rand von Y anschmiegt. Nach den Überlegung von **4** genügt es, wenn man $\rho(y) \to 0$ für $y \to \infty$ unter der Nebenbedingung $N(y) = 1$ nachweist.

Man überlegt sich dazu zuerst, dass für $N(y_\nu) = 1$ die beiden Aussagen $y_\nu \to \infty$ und $y^*_\nu \to \infty$ gleichbedeutend sind. Ist nämlich zum Beispiel $y_\nu \to \infty$ und würde es eine beschränkte konvergente Teilfolge der y^*_ν geben, dann können wir ohne Einschränkung $y^*_\nu \to a$ annehmen und a ist dann sicher ein Randpunkt von Y. Nach Lemma 5 strebt $N(y^*_\nu)$ gegen Null und $N(y_\nu)N(y^*_\nu) = 1$ ergibt einen Widerspruch.

Wendet man Lemma 1 auf die in **7** bewiesene Gleichung $y'Sy^* = n$ an, so folgt $\rho(y) | y^* | \leqq n$ für alle y aus Y, d.h. $\rho(y) \to 0$ falls $y^* \to \infty$. Wir hatten ober oben schon gesehen, dass für $y \to \infty$, $N(y) = 1$, auch $y^* \to \infty$ gilt, so dass also

$$\rho(y) \to 0 \text{ falls } y \to \infty, N(y) = 1$$

richtig ist.

10. Die Gamma-Funktion eines Positivitätsbereiches. Wie bisher sei Y ein homogener Positivitätsbereich. In Verallgemeinerung der Euler'schen Definition der Γ-Funktion definieren wir für eine komplexe Variable s mit $\operatorname{Re} s \geqq 1$ und a aus Y die *Gamma-Funktion* von Y durch

$$\Gamma(Y, s) = [N(a)]^s \int_Y e^{-a'Sy} [N(y)]^{s-1} \, dy.$$

Zum Nachweis der Konvergenz dieses Integrals überlegt man sich, dass $N(y) \leqq \rho | y |^n$ mit geeignetem $\rho > 0$ für alle y aus Y gilt. Lemma 1 zeigt dann, dass $\Gamma(Y, s)$ für $\operatorname{Re} s \geqq 1$ absolut konvergiert und dort eine holomorphe Funktion von s darstellt. Wegen der Homogenität von Y ist sofort zu sehen, dass die rechte Seite nicht von der Wahl des Punktes a abhängt.

Eine genauere Untersuchung dieser Gamma-Funktion wäre von Interesse, denn man kann $\Gamma(Y, s)$ als einfachste analytische Invariante von Y auffassen. Es ist zu vermuten, dass $\Gamma(Y, s)$ sich bis auf elementare Faktoren als Produkt von Funktionen der Form $\Gamma(\alpha s + \beta)$ schreiben lässt. Für eine einfache Klasse von Bereichen wurde dies von S. Bochner bewiesen.

Wir wollen jetzt eine wichtige Klasse von Integralen über Y berechnen, d.h. auf $\Gamma(Y, s)$ zurückführen.

LEMMA 7. *Ist $f(x)$ für positive x stückweise stetig und existiert für $\operatorname{Re} s \geqq 1$ das Integral*

$$I(f; s) = \int_0^\infty f(x) x^{ns-1} \, dx$$

absolut, dann gilt für diese s

$$\int_Y f(a'Sy)[N(y)]^{s-1}dy = \frac{I(f;s)\Gamma(Y,s)}{\Gamma(ns)[N(a)]^s}.$$

Beweis: Es sei $d\omega(x)$ das durch dy auf der Hyperfläche $|y| = x$ induzierte Flächenelement, d.h. $d\omega(x) = x^{n-1}d\omega$, wobei dann $d\omega$ das Flächenelement für $|y| = 1$ ist. Jetzt hat man für das zu bestimmende Integral

$$\int_0^\infty dx \int_{|y|=1} d\omega\, x^{n-1} f(x\cdot a'Sy)[N(xy)]^{s-1}$$

$$= \int_0^\infty f(x)x^{ns-1}dx \int_{|y|=1}[N(y)]^{s-1}(a'Sy)^{-ns}d\omega = I(f;s)\cdot I(a),$$

wobei also das zweite Integral mit $I(a)$ bezeichnet ist. Zur Bestimmung $I(a)$ setzt man $f(x) = e^{-x}$ und erhält

$$[N(a)]^{-s}\Gamma(Y,s) = \Gamma(ns)\cdot I(a)$$

also die Behauptung.

Eine weitere Integralbeziehung geben wir ohne Beweis an: Es gilt für $\operatorname{Re} s \geqq 1$

$$\int_Y \frac{dy}{[N(a+y)]^{s+1}} = \frac{\omega}{[N(a)]^s}\,\frac{\Gamma(Y,s)}{\Gamma(Y,s+1)}$$

und das linke Integral konvergiert absolut. Dabei bedeutet ω die zur Normierung von $N(y)$ in **5** benutzte Konstante.

11. Beispiele von Positivitätsbereichen.

a) *Der triviale Bereich.* Es sei Y^n die Menge der Vektoren des R^n mit positiven Komponenten und S eine beliebige symmetrische Permutationsmatrix von n Zeilen, d.h. eine symmetrische Matrix, bei der in jeder Zeile und jeder Spalte genau eine 1 und sonst Nullen stehen. Offenbar ist Y^n ein nicht ausgearteter Positivitätsbereich mit S als Charakteristik, den wir auch den *trivialen Bereich* des R^n nennen wollen. Schon an diesem einfachen Beispiel sieht man, dass die Charakteristik S keineswegs immer positiv definit ist, unter den symmetrischen Permutationsmatrizen kommen vielmehr alle möglichen Signaturen vor.

Nach **5** berechnet man die Norm bis auf einen konstanten Faktor zu

$$N(y) = \prod_k y^k \text{ und } ds^2 = \sum_k (dy^k/y\ )^2.$$

Für die Involution nach **7** findet man $y^{*\prime} = S^{-1}(1/y^1, 1/y^2, \cdots, 1/y^n)$. Ohne Schwierigkeiten zeigt man ferner, dass die Automorphismengruppe $\Sigma(Y^n)$

genau aus den Produkten PD einer Permutationsmatrix P und einer Diagonalmatrix D mit positiven Diagonalelementen besteht. Die Gamma-Funktion von Y^n erhält man in der Form $\Gamma(Y^n, s) = [\Gamma(s)]^n$.

b) *Positive quadratische Formen.* Sei m eine natürliche Zahl und $n = m(m+1)/2$. Für jede symmmetrische m-reihige reelle Matrix $\mathfrak{Y} = (y_{kl})$ setzen wir

$$f(\mathfrak{Y}) = (y_{11}, y_{22}, \cdots, y_{mm}, y_{12}, y_{13}, \cdots, y_{m-1,m})'.$$

Die Abbildung $\mathfrak{Y} \to f(\mathfrak{Y})$ ist dann eine eineindeutige Abbildung des Raumes aller symmetrischen reellen Matrizen von m Zeilen auf den R^n. Sei nun

$$Y_P{}^{(m)} = \{y \colon y = f(\mathfrak{Y}), \mathfrak{Y}^{(m)} \text{ positiv definit}\}$$

$$S = \begin{bmatrix} 1 & & & & & & \\ & \cdot & & & & & \\ & & \cdot & & & 0 & \\ & & & 1 & & & \\ & & & & 2 & & \\ & & & & & \cdot & \\ & & & & & & \cdot \\ & 0 & & & & & \cdot \\ & & & & & & & 2 \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}1\\1\\1\end{matrix}}\right\} m \\ \left.\vphantom{\begin{matrix}1\\1\\1\\1\\1\end{matrix}}\right\} n-m \end{matrix}$$

dann prüft man leicht, dass für $a = f(\mathfrak{A})$ und $y = f(\mathfrak{Y})$ stets

$$a'Sy = Spur(\mathfrak{A}\mathfrak{Y})$$

erfüllt ist. Offenbar ist dann $Y_P{}^{(m)}$ ein nicht ausgearteter Positivitätsbereich des R^n.

Ist $\mathfrak{R}$ eine reelle nicht ausgeartete Matrix von m Zeilen, so liefert $\mathfrak{Y} \to \mathfrak{R}'\mathfrak{Y}\mathfrak{R}$ eine eineindeutige Abbildung der positiv definiten Matrizen auf sich und in

$$f(\mathfrak{R}'\mathfrak{Y}\mathfrak{R}) = Wf(\mathfrak{Y}), \qquad W = W^{(n)} = W(\mathfrak{R})$$

hängt die Matrix W nur von $\mathfrak{R}$ ab und ihre Komponenten sind homogene quadratische Polynome in den Elementen von $\mathfrak{R}$. Die Gruppe Σ_0 der Matrizen $W = W(\mathfrak{R})$, $|\mathfrak{R}| \neq 0$, ist eine Untergruppe von $\Sigma(Y_P{}^{(m)})$, es ist leicht zu zeigen, dass Σ_0 mit der Automorphismengruppe übereinstimmt. Man sieht aber sofort, dass schon Σ_0 auf $Y_P{}^{(m)}$ transitiv wirkt, d.h. der Positivitätsbereich homogen ist. Ohne Schwierigkeiten findet man

$$N(y) = |\mathfrak{Y}|^{(m+1)/2}, \qquad \text{falls } y = f(\mathfrak{Y}),$$

$$ds^2 = (m+1)/2\, Spur(\mathfrak{Y}^{-1}\delta\mathfrak{Y}\mathfrak{Y}^{-1}\delta\mathfrak{Y}), \qquad \delta\mathfrak{Y} = (dy_{kl}),$$

$$y^* = (m+1)/2\, f(\mathfrak{Y}^{-1}), \qquad \text{falls } y = f(\mathfrak{Y})$$

und nach C. L. Siegel ist

$$\Gamma(Y_P^{(m)}, s) = \pi^{m(m-1)/4}\prod_{k=0}^{m-1}\Gamma(\tfrac{1}{2}(s[m+1]-k)).$$

c) *Positive hermitesche Formen.* Man setzt hier $n = m^2$ und betrachtet die hemiteschen Matrizen von m Zeilen und Spalten

$$\mathfrak{Z} = \mathfrak{Z}' = (z_{kl}) = \mathfrak{Y} + i\mathfrak{X}, \quad \mathfrak{Y}' = \mathfrak{Y}, \quad \mathfrak{X}' = -\mathfrak{X}.$$

Für schiefsymmetrische Matrizen $\mathfrak{X} = (x_{kl})$ setzen wir analag zu b)

$$g(\mathfrak{X}) = (x_{12}, x_{13}, \cdots, x_{m-1,m})'$$

und definieren in der dortigen Bezeichnung

$$h(\mathfrak{Z}) = \begin{pmatrix} f(\mathfrak{Y}) \\ g(\mathfrak{X}) \end{pmatrix}, \qquad \mathfrak{Z} = \mathfrak{Y} + i\mathfrak{X}.$$

Wie in b) findet man

$$a'Sz = Spur(\mathfrak{A}\mathfrak{Z}), \quad a = h(\mathfrak{A}), \quad z = h(\mathfrak{Z})$$

und weisst nach, dass die Menge $y = h(\mathfrak{Z})$, für die $\mathfrak{Z}$ eine hermitesch positiv definite Matrix ist, einen nicht ausgearteten Positivitätsbereich $Y_P^{(m)}$ bilden. Eine hermitesche Matrix $\mathfrak{Z}$ heisst dabei *hermitesch positiv definit,* wenn $\mathfrak{x}'\mathfrak{Z}\mathfrak{x}$ für alle komplexen Vektoren $\mathfrak{x} \neq 0$ positiv ausfällt. Die Norm berechnet man hier zu

$$N(y) = \| \mathfrak{Z} \|^m \text{ falls } y = h(\mathfrak{Z})$$

und nach H. Braun gilt

$$\Gamma(Y_H^{(m)}, s) = (2\pi)^{m(m-1)/2}\prod_{k=0}^{m-1}\Gamma(ms+k)$$

d) *Positivitätsbereiche im R^3.* Von einem ganz anderen Typus als die bisherigen Beispiele sind die Bereiche, die wir jetzt untersuchen wollen. Sei dazu

$$0 < \rho < 1, \qquad \alpha = 1/\rho(\rho/(1-\rho))^{1-\rho}, \qquad S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \alpha \end{pmatrix}.$$

Das Ziel dieses letzten Abschnittes ist es, zu zeigen, dass der Bereich

$$Y(\rho) = \{y : y = \begin{pmatrix} u \\ v \\ w \end{pmatrix}, u > 0, v > 0, |w| < u^\rho v^{1-\rho}\}$$

ein Positivitätsbereich des R^3 mit S als Charakteristik ist.

Die Komponenten der Punkte des R^3 bezeichnen wir stets mit u, v und w. Sind jetzt y_1 und y_2 zwei Punkte aus $Y(\rho)$, dann hat man

$$y'_1 S y_2 = u_1u_2 + v_1v_2 + \alpha w_1 w_2, \quad |w_j| < u_j^{\rho} v_j^{1-\rho}$$
$$> u_1u_2 + v_1v_2 - \alpha(u_1u_2)^{\rho}(v_1v_2)^{1-\rho} = u_1u_2 p(v_1v_2/u_1u_2)$$

mit $p(t) = 1 + t - \alpha t^{1-\rho}$. Man überlegt sich sofort, dass $\tau = (1-\rho)/\rho$ das einzige Minimum von $p(t)$ ist. Wegen $p(\tau) = 0$ ist dann aber $p(t) \geqq 0$ für alle $t \geqq 0$. Damit ist (P. 1) nachgewiesen. Ist y_1 dagegen ein Punkt des Komplementes von $Y(\rho)$, dann ist (P. 2) in den Fällen $u_1 \leqq 0$ oder $v_1 \leqq 0$ leicht zu sehen. Sind dagegen u_1 und v_1 beide positiv, dann ist $|w_1| \geqq u_1^{\rho} v_1^{1-\rho}$. Wir wählen jetzt y_2 aus $\bar{Y}(\rho)$ mit den Nebenbedingungen

$$v_1v_2/u_1u_2 = \tau, \quad w_2 = -(w_1/|w_1|)u_2^{\rho}v_2^{1-\rho}$$

und erhalten

$$y'_1 S y_2 = u_1u_2 + v_1v_2 + \alpha w_1 w_2$$
$$= u_1u_2 p(\tau) - \alpha(|w_1| - u_1^{\rho}v_1^{1-\rho})u_2^{\rho}v_2^{1-\rho} \leqq 0,$$

also ebenfalls (P. 2).

Die Bereiche $Y(\rho)$ sind für $2\rho \neq 1$ bisher nicht untersucht, der Fall $2\rho = 1$ führt auf die positiven quadratischen Formen in zwei Variablen. Da aber alle $Y(\rho)$ gemeinsam mit den in a), b) und c) aufgeführten wichtigen Beispielen als Positivitätsbereiche unter einem gemeinsamen Gesichtspunkt erscheinen, wäre eine Untersuchung der $Y(\rho)$ von grossem Interesse. Es ist zu vermuten, dass die $Y(\rho)$, $2\rho \neq 1$, nicht mehr homogen sind.

Man kann schliesslich noch ohne grosse Schwierigkeiten zeigen, dass die Bereiche

$$Y^3, \; Y_P^{(2)} \text{ und } Y(\rho) \qquad (0 < \rho < \tfrac{1}{2})$$

bis auf affine Abbildungen *alle* Positivitätsbereiche des R^3 liefern. Eine Klassifizierung der Positivitätsbereiche ist auch für $n = 4$ noch relativ leicht möglich.

UNIVERSITÄT MÜNSTER/WESTF.

ON A PROBLEM OF WEYL IN THE THEORY OF SINGULAR STURM-LIOUVILLE EQUATIONS.* [1]

By N. Aronszajn.

1. In his famous paper [7] Weyl proposed the following problem: Consider the singular Sturm-Liouville equation

$$(1.1) \qquad -(px')' + (q-\zeta)x = 0, \qquad 0 \leqq t < \infty,$$

with boundary condition at the regular endpoint $t = 0$

$$(1.2_\alpha) \qquad \sin\alpha\, x(0) - \cos\alpha\, p(0)x'(0) = 0.$$

Suppose that the equation is in the limit-point case. Is the continuous spectrum preserved if the boundary condition (1.2_α) is changed into another, (1.2_β)?

A. Wintner stated a wider problem in a more general framework, namely, to investigate the changes in the spectrum of a self-adjoint problem when it is submitted to a one-dimensional perturbation (see P. Hartman [3]). Apparently very few papers have been published on this subject.[2] There is one by Putnam [5] which deals with the problem in its original setting.

Using a general theory of finite dimensional perturbations of spectral problems [3] the author was able to investigate the general problem of Wintner, showing in particular, that a one-dimensional perturbation can transform a pure point spectrum into a pure continuous spectrum and that for any self-adjoint spectral problem with simple spectrum, the part of the pure point spectrum contained in the limit spectrum can be made to disappear by a suitable one-dimensional perturbation. It was also proved that a part of the continuous spectrum, the *absolutely continuous spectrum,* is always preserved in finite dimensional perturbations. After these results were obtained, it became apparent that in the original setting of a singular Sturm-Liouville problem, most of these results could be achieved by using only the classical

* Received January 21, 1957.

[1] Written under contract with the Office of Naval Research, Contract Nonr 583(04).

[2] The author's attention was drawn to this problem in a conversation with P. Hartman.

[3] This theory will be presented in a paper by W. F. Donoghue and the author on "Finite dimensional perturbations of spectral problems and approximation methods."

background and recent results of Gelfand and Levitan [2]. This treatment of Weyl's problem in its original setting is the subject of the present paper.

In Section 2 we recall some properties of measures on the real axis and their relations to analytic functions in the upper half plane with positive imaginary part. In Section 3 we recall some essential facts from Weyl's theory of singular Sturm-Liouville equations and apply the developments of Section 2 to obtain our basic theorem concerning the conservation of the absolutely continuous spectrum (Theorem 1). In Section 4 we formulate an essential part of the basic result of Gelfand-Levitan [2] in a slightly improved form due to M. Krein [4]. From this result we deduce a sufficient condition for a measure μ to be equivalent to the spectral measure of a Sturm-Liouville problem in the limit-point case, with $p(t) \equiv 1$ (Theorem 2). This sufficient condition is used in the next sections to construct counter examples for Weyl's problem. In Section 5 we construct a Sturm-Liouville equation which for one boundary condition has a pure continuous singular spectrum concentrated on a closed set of Lebesgue measure 0 whereas for all other boundary conditions it has a pure point spectrum with all eigenvalues isolated. In Section 6 we give first a characterization of the point spectrum for a boundary condition (1.2_β) in terms of the spectral measure corresponding to another boundary condition, (1.2_α) (Theorem 4). Using this characterization we construct an example of a Sturm-Liouville problem which for one boundary condition has a pure point spectrum dense on the whole real axis, whereas for all other boundary conditions the spectrum is purely continuous.

2. We shall consider Lebesgue-Stieltjes measures on the real axis R; they will be given by a nondecreasing function $\mu(\lambda)$ normalized as follows: $\mu(0) = 0$, $\mu(\lambda) = \frac{1}{2}[\mu(\lambda+) + \mu(\lambda-)]$. The letters μ, ν, ρ, etc. will stand for the functions as well as the corresponding measures: $\mu(\lambda)$ is the value of the function at λ; $\mu(S)$ is the measure of the set S; $\mu\{\lambda\}$ is the measure of the set $\{\lambda\}$ composed of the single point λ, etc. The Lebesgue measure of a set S will be written $|S|$.

All notions relative to a measure when written without mention of the measure mean relative to the Lebesgue measure; for instance: a.e. (almost everywhere), absolutely continuous, etc.

A *support* of μ is a set S such that $\mu(R - S) = 0$. We recall that μ and ν are called *orthogonal* if some of their supports are disjoint.

The *closed support* of μ is the smallest among all closed sets which are supports of μ. Two equivalent measures have the same closed support;

the converse is not generally true. In our considerations the closed supports will not be the most important; we will not even require that a support be contained in the closed support. However, we will impose another restriction on supports (see (2.1) below).

A support S of μ is said to be *minimal rel.* ν if for any smaller support $S_1 \subset S$ of μ, $\nu(S - S_1) = 0$. It is clear that each support S of μ is minimal rel. μ and that two measures μ and ν are equivalent if and only if they have a common support which is minimal for each one of them rel. the other.

In accordance with our general convention we say that a support of μ is minimal if it is minimal rel. to the Lebesgue measure. From now on we will assume

(2.1) *All supports of measures will be assumed minimal.*

It follows that if S is a support of a singular measure, $|S| = 0$. Also if two absolutely continuous measures have the same support, they are equivalent.

For a point measure (i.e. a measure with an enumerable support) we will consider only the smallest support (i.e. the set of all points with positive measure).

Corresponding to a decomposition of μ into mutually orthogonal measures, ν_k, $\mu = \sum \nu_k$, we will consider supports S_k of ν_k as *admissible* if they are mutually disjoint. We are especially interested in the Lebesgue-Jordan (L.-J.) decompositions $\mu = \mu^0 + \mu^1$, $\mu^1 = \mu' + \mu''$, where μ^0 is the absolutely continuous part, μ^1 the singular part, μ' the singular continuous part, and μ'' the point measure corresponding to μ. Corresponding admissible supports for these decompositions can be obtained from a classical theorem of De la Vallée Poussin (see e.g. Saks [6]):

$$S^0{}_P = \underset{\xi}{\mathrm{E}}[d\mu(\xi)/d\lambda \textit{ exists and } 0 < d\mu(\xi)/d\lambda < \infty],$$

$$S^1{}_P = \underset{\xi}{\mathrm{E}}[d\mu(\xi)/d\lambda \textit{ exists and } d\mu(\xi)/d\lambda = \infty],$$

$$S''{}_P = S'' = \underset{\xi}{\mathrm{E}}[\mu(\lambda) \textit{ is discontinuous at } \xi], \qquad S'{}_P = S^1{}_P - S''{}_P.$$

For our needs it will be more convenient to replace these supports by a slightly different system connected with analytic functions with positive imaginary part in the half-plane, $\operatorname{Im} \zeta > 0$. We recall some of the fundamental properties of these functions. Each such function $\phi(\zeta)$ has a canonical representation of the form

$$\phi(\zeta) = \omega\zeta + \sigma + \int_{-\infty}^{\infty} [(\lambda - \zeta)^{-1} - \lambda(\lambda^2 + 1)^{-1}] d\mu(\lambda), \tag{2.2}$$

with $\omega \geqq 0$, σ real, and μ a positive measure such that

$$\int (\lambda^2+1)^{-1} d\mu < \infty. \tag{2.3}$$

Conversely, each choice of such ω, σ, and μ, defines by (2.2) a function $\phi(\zeta)$ with positive imaginary part for $\operatorname{Im}\zeta > 0$.

In the following properties we will use for a real ξ the expression "$\zeta \to \xi$ in an angle" meaning that for some ϵ, $0 < \epsilon < \frac{1}{2}\pi$, ζ converges to ξ remaining inside the angle $\epsilon < \operatorname{Arg}(\zeta - \xi) < \pi - \epsilon$. The statements below are well known.

I. *For ξ a.e. in R, $\phi(\zeta)$ converges to a finite limit when $\zeta \to \xi$ in any angle.*

II. *If $d\mu(\xi)/d\lambda$ exists (may be finite or infinite)* $\operatorname{Im}\phi(\lambda) \to \pi(d\mu(\xi)/d\lambda)$ *when $\zeta \to \xi$ in any angle.*

III. *If the upper derivative of μ at ξ is infinite, then* $\limsup_{\eta \searrow 0} \operatorname{Im}\phi(\xi + i\eta) = \infty$.

IV. *For every ξ, $(\xi - \zeta)\phi(\zeta)$ converges to $\mu\{\xi\} = \mu(\xi+) - \mu(\xi-)$ when $\zeta \to \xi$ in any angle.*

We now define our *standard* supports for the parts of μ in L.-J. decomposition as follows: We suppose that μ satisfies (2.3) and with any $\omega \geqq 0$ and σ real, we form the function $\phi(\zeta)$ by (2.2). We then put

$$S^0 = \mathop{\mathrm{E}}_{\xi}[\lim \phi(\zeta) \text{ exists and is finite and } \lim \operatorname{Im}\phi(\zeta) > 0 \text{ when } \zeta \to \xi \text{ in any angle}]. \tag{2.4^0}$$

$$S^1 = \mathop{\mathrm{E}}_{\xi}[\operatorname{Im}\phi(\zeta) \to \infty \text{ when } \zeta \to \xi \text{ in any angle}]. \tag{2.4^1}$$

$$S' = \mathop{\mathrm{E}}_{\xi}[\operatorname{Im}\phi(\zeta) \to \infty \text{ and } (\zeta - \xi)\phi(\zeta) \to 0 \text{ when } \zeta \to \xi \text{ in any angle}]. \tag{2.4$'$}$$

$$S'' = \mathop{\mathrm{E}}_{\xi}[\lim (\xi - \zeta)\phi(\zeta) > 0 \text{ when } \zeta \to \xi \text{ in any angle}]. \tag{2.4$''$}$$

That these are admissible supports follows immediately from the statements above by using the supports of de la Vallée Poussin. We have here $S^1 \supset S^1{}_P$, whereas there is no generally valid inclusion relation between S^0 and $S^0{}_P$.

It is obvious that the standard supports depend only on the measure and not on the choice of the constants ω and σ. It would be easy to define them without assuming the restriction (2.3) on μ but it is immaterial for us, since all the measures in which we are interested satisfy (2.3).

All the standard supports are contained in the closed support of μ but they are not necessarily contained in the closed supports of the corresponding parts of μ. It may even happen that the singular measure μ^1 is $=0$ with the support S^1 nonvanishing.

Remark. For two equivalent measures μ and ν the standard supports may be quite different. It can be proved quite easily, however, that if the derivative $d\mu(\lambda)/d\nu$ is given by a function $\tau(\lambda)$, continuous and positive on the whole real axis, and satisfying on each finite interval a Lipschitz condition, the standard supports for μ and ν are the same.

3. In the original setting for Weyl's treatment of equation (1.1) with boundary condition (1.2_α), it is assumed that $p(t)$ and $q(t)$ are continuous and $p(t)$ is positive for $0 \leqq t < \infty$. As usual, we form the solutions $\phi_\alpha(t,\zeta)$ and $\psi_\alpha(t,\zeta)$ of equation (1.1) determined by the initial values

$$(3.1)\qquad \begin{aligned} \phi_\alpha(0,\zeta) &= \sin\alpha, & \phi_\alpha'(0,\zeta)p(0) &= -\cos\alpha, \\ \psi_\alpha(0,\zeta) &= \cos\alpha, & \psi_\alpha'(0,\zeta)p(0) &= \sin\alpha.\text{[4]} \end{aligned}$$

In the limit-point case there exists for each non-real ζ a unique complex number $m_\alpha(\zeta)$—the limit point—such that $\phi_\alpha(t,\zeta)+m_\alpha(\zeta)\psi_\alpha(t,\zeta)$ is in $L^2(0,\infty)$. It is known that $m_\alpha(\zeta)$ is analytic and with positive imaginary part in the half-plane $\operatorname{Im}\zeta>0$. It is also proved that in the canonical representation of $m_\alpha(\zeta)$ of type (2.2), the constant $\omega=0$. Hence

$$(3.2)\qquad m_\alpha(\zeta) = \sigma_\alpha + \int_{-\infty}^{\infty} [(\lambda-\zeta)^{-1} - \lambda(\lambda^2+1)^{-1}]\,d\rho_\alpha(\lambda),$$

where the measure ρ_α satisfies condition (2.3).

The measure ρ_α is the Weyl spectral measure of the problem. Introducing the L.-J. decompositions of ρ_α, we get the traditional spectra defined as follows: The closed support of ρ_α is the spectrum of our problem; the support of ρ''_α is the point spectrum; the closed support of $\rho^0_\alpha+\rho'_\alpha$ is the continuous spectrum. The closed support of ρ_α minus all its isolated points (which belong to the point spectrum) forms the limit spectrum (sometimes called the essential spectrum). We may add here the absolutely continuous spectrum which is the closed support of ρ^0_α, and the singular continuous spectrum which is the closed support of ρ'_α.[5]

---

[4] The primes denote differentiation with respect to t.

[5] It is well known that these sets by themselves do not give an adequate representation of the spectral decomposition corresponding to the problem. It is rather the corresponding measures (or any measures equivalent to them) which should be considered as representing the spectra.

The function $m_\alpha(\zeta)$ and $m_\beta(\zeta)$ for two different boundary conditions satisfy the following relation, where we put $\gamma = \beta - \alpha$:

$$m_\beta(\zeta) = \{m_\alpha(\zeta)\operatorname{cotg}\gamma - 1\}/\{m_\alpha(\zeta) + \operatorname{cotg}\gamma\}.^6$$

It will be more convenient for us to write this formula as follows:

$$(3.3) \qquad [m_\alpha(\zeta) + \operatorname{cotg}\gamma][m_\beta(\zeta) - \operatorname{cotg}\gamma] = -1/\sin^2\gamma.^7$$

We may define the standard supports $S^0{}_\alpha, S^1{}_\alpha, \cdots, S^0{}_\beta, S^1{}_\beta, \cdots$, for the L.-J. decompositions of ρ_α and ρ_β by using, for $\phi(\zeta)$, the functions $m_\alpha(\zeta) + \operatorname{cotg}\gamma$ and $m_\beta(\zeta) - \operatorname{cotg}\gamma$ respectively. A simple inspection of the definitions (2.4^0) and (2.4^1) and formula (3.3) shows that $S^0{}_\alpha = S^0{}_\beta$ and that $S^1{}_\alpha \cap S^1{}_\beta = 0$. Hence our theorem:

THEOREM 1. *For two distinct boundary conditions* (1.2_α) *and* (1.2_β) *the absolutely continuous parts of the corresponding spectral measures are equivalent, whereas the singular parts are orthogonal.*

This gives the positive part of our answer to Weyl's problem. The absolutely continuous spectrum does not change when the boundary condition is changed.

In the following sections we will construct examples showing that in general the singular continuous spectrum is not preserved.

Remark. In his paper, Putnam [5] proves the following statement: If in some open interval I, ρ_α is absolutely continuous relative to ρ_β for one $\beta \neq \alpha$, then it is absolutely continuous in I relative to $\rho_{\beta'}$ for all β'. From our theorem we can conclude more, namely: that under the hypothesis of Putnam, ρ_α in I is absolutely continuous and equivalent to the absolutely continuous part of $\rho_{\beta'}$ for all β'.

4. In order to construct the examples mentioned above, we need some information about sufficient conditions for a measure μ to be equivalent to a spectral measure of a singular Sturm-Liouville problem in the limit-point

---

[6] This formula can be found in the book of Coddington and Levinson [1], Exercise 8, page 257.

[7] In the general theory of perturbations of spectral problems (developed in the forthcoming paper of W. F. Donoghue and the author) where the change from one boundary condition to another takes the aspect of a one-dimensional perturbation, the two square brackets multiplied by $\pm \sin\gamma$ are the determinants of the two opposite perturbations, and formula (3.3) is a special case of the general composition formula for the determinants of perturbations.

case. For this purpose we will recall briefly some recent results of Gelfand-Levitan [2] in an improved form due to Krein [4].

For a measure μ define the function $\sigma(\lambda)$ for $\lambda \geqq 0$

(4.1) $$\sigma(\lambda) = \mu(\lambda) - 2\lambda^{\frac{1}{2}}/\pi, \text{ for } \lambda \geqq 0.$$

Consider now the following conditions:

(4.2) *The closed support of μ possesses a finite limit point.*

(4.3) *For some $\theta < 2$, $\int_{-\infty}^{0} \exp\{(-\lambda)^{\frac{1}{2}}t\} d\mu(\lambda) < \exp(t^{\theta})$ for sufficiently large positive t.*

(4.4_k) *The integral $\int_{1}^{\infty} \lambda^{-1}\cos(\lambda^{\frac{1}{2}}t)\, d\sigma(\lambda)$ is absolutely convergent for all $0 \leqq t < \infty$ and defines a function $a(t)$ with continuous $k+3$ derivatives.*

THEOREM OF GELFAND-LEVITAN. *If $\mu(\lambda)$ satisfies the conditions (4.2), (4.3), and (4.4_k) for some $k = 0, 1, 2, \cdots$ then μ is the Weyl spectral measure ρ_α for a uniquely determined Sturm-Liouville problem $-x'' + (q - \zeta)x = 0$, $x'(0) = \tan\alpha\, x(0)$ where $-\frac{1}{2}\pi < \alpha < \frac{1}{2}\pi$, $q(t)$ has k continuous derivatives and the problem is in the limit-point case.*

Remark 1. From the relation between $a(t)$ and $q(t)$ (given by Gelfand and Levitan [2]) it follows easily that if $a(t)$ is analytic in the real variable t for $0 \leqq t < \infty$, then so is also the function $q(t)$.

We can now give sufficient conditions for a measure μ to be equivalent to a spectral measure of a singular Sturm-Liouville problem in the limit point case.

THEOREM 2. *Let μ be a measure satisfying (4.2) and let $\{\xi_l\}$, $l = 0, 1, 2, \cdots$, be a strictly increasing sequence such that $\xi_0 = 0$, $\xi_1 \geqq 1$, $\xi_l \to \infty$, $\mu\{\xi_l\} = 0$ for $l \geqq 1$, and $\mu(\xi_{l+1}) - \mu(\xi_l) > 0$ for all l. If for some $k \geqq 0$, the condition*

(4.5) $$\sum_{l=1}^{\infty} \xi_l^{\frac{1}{2}(k-1)} (\xi_{l+1} - \xi_l)^2 < \infty$$

is satisfied, there exists a measure ρ equivalent to μ such that ρ is the spectral measure ρ_α of a Sturm-Liouville problem in the limit-point case, with $p(t) = 1$, $q(t)$ k times continuously differentiable and $-\frac{1}{2}\pi < \alpha < \frac{1}{2}\pi$. In addition, for each bounded part of a support of μ there exist constants C_0 and C_1 such that $0 < C_0 < d\rho/d\mu < C_1 < \infty$.

Proof. We construct a measure ρ equivalent to μ and which satifies the conditions in the Gelfand-Levitan theorem. Obviously, for every ρ equivalent to μ (4.2) is satisfied. In order to fulfill condition (4.3) we proceed as follows. In each open interval $(-n;\, -n+1)$, for $n = 1, 2, \cdots$ we put $d\rho(\lambda) = \exp(-\lambda^2)[\mu(-n+1) - \mu(-n)]^{-1} d\mu(\lambda)$ unless there is no measure in this interval, in which case we put $d\rho(\lambda) = 0$. For each $n = 1, 2, \cdots$, for which $\mu\{-n\} > 0$, we put $\rho\{-n\} = \exp(-n^2)$; if $\mu\{-n\} = 0$, we put $\rho\{-n\} = 0$. It is easy to check that ρ so defined on the left axis is equivalent there to μ, satisfies condition (4.3), and also the last condition in the theorem.

We define now the measure ρ on the half axis $\lambda \geqq 0$ by putting $\rho\{0\} = [2\xi_1^{\frac{1}{2}}/\pi\mu(\xi_1)]\mu\{0\}$ and on each interval

$$(\xi_l;\, \xi_{l+1}] : d\rho(\lambda) = 2\pi^{-1}[\xi_{l+1}^{\frac{1}{2}} - \xi_l^{\frac{1}{2}}][\mu(\xi_{l+1}) - \mu(\xi_l)]^{-1} d\mu(\lambda).$$

$\rho(\lambda)$ defined in this way on the positive axis is obviously equivalent to μ and satisfies the last condition in the theorem. Also like $\mu(\lambda)$, $\rho(\lambda)$ is continuous at the points ξ_l, $l \geqq 1$.

To prove condition (4.4_k) consider for $\lambda \geqq 0$, $\sigma(\lambda) = \rho(\lambda) - 2\pi^{-1}\lambda^{\frac{1}{2}}$. Since $\rho(\xi_0) = \rho(0) = 0$ and $\rho(\xi_{l+1}) - \rho(\xi_l) = 2\pi^{-1}[\xi_{l+1}^{\frac{1}{2}} - \xi_l^{\frac{1}{2}}]$ we have $\sigma(\xi_l) = 0$ for all l.

We show first that the integral defining $a(t)$ is absolutely convergent. In fact

$$\int_1^\infty \lambda^{-1} \,|\, d\sigma(\lambda) | = \int_1^{\xi_1} \lambda^{-1} \,|\, d\sigma(\lambda) | + \sum_{l=1}^\infty \int_{\xi_l}^{\xi_{l+1}} \lambda^{-1} \,|\, d\sigma(\lambda) |,$$

$$\int_{\xi_l}^{\xi_{l+1}} \lambda^{-1} \,|\, d\sigma(\lambda) | \leqq 2\pi^{-1} \int_{\xi_l}^{\xi_{l+1}} \lambda^{-1} d\lambda^{\frac{1}{2}} + 2\pi^{-1}[\xi_{l+1}^{\frac{1}{2}} - \xi_l^{\frac{1}{2}}][\mu(\xi_{l+1}) - \mu(\xi_l)]^{-1}$$

$$\times \int_{\xi_l}^{\xi_{l+1}} \lambda^{-1} d\mu(\lambda) \leqq 2\pi^{-1} \left\{ \int_{\xi_l}^{\xi_{l+1}} \lambda^{-1} d\lambda^{\frac{1}{2}} + \xi_l^{-1}[\xi_{l+1}^{\frac{1}{2}} - \xi_l^{\frac{1}{2}}] \right\}.$$

Condition (4.5) implies that there exists a constant c such that $\xi_{l+1} < c\xi_l$, $l \geqq 1$. It follows that the last expression is majorated by $2(1+c)\pi^{-1} \int_{\xi_l}^{\xi_{l+1}} \lambda^{-1} d\lambda^{\frac{1}{2}}$ and hence

$$\int_1^\infty \lambda^{-1} \,|\, d\sigma(\lambda) | \leqq \int_1^{\xi_1} \lambda^{-1} \,|\, d\sigma(\lambda) | + 2(1+c)\pi^{-1} \int_{\xi_1}^\infty \lambda^{-1} d\lambda^{\frac{1}{2}} < \infty.$$

Next we express $a(t)$ as follows

$$a(t) = \int_1^{\xi_1} \lambda^{-1} \cos(\lambda^{\frac{1}{2}} t)\, d\sigma(\lambda) + \sum_{l=1}^\infty \int_{\xi_l}^{\xi_{l+1}} \lambda^{-1} \cos(\lambda^{\frac{1}{2}} t)\, d\sigma(\lambda) \equiv u_0(t) + \sum_{l=1}^\infty u_l(t).$$

To prove condition (4.4_k) it is now sufficient to show that the $(k+3)$-th derivatives of $u_l(t)$ form an absolutely and uniformly convergent series in every finite interval of the variable t.

By partial integration and using the fact that $\sigma(\xi_l)=0$, we get

$$|u_l^{(k+3)}(t)| = \left| \int_{\xi_l}^{\xi_{l+1}} \lambda^{\frac12(k+1)} \begin{matrix}\cos\\ \sin\end{matrix} (\lambda^{\frac12}t)\, d\sigma(\lambda) \right|$$
$$= \left| \int_{\xi_l}^{\xi_{l+1}} [\tfrac12(k+1)\lambda^{\frac12(k-1)} \begin{matrix}\cos\\ \sin\end{matrix} (\lambda^{\frac12}t) \mp \tfrac12 t\lambda^{\frac12 k} \begin{matrix}\sin\\ \cos\end{matrix} (\lambda^{\frac12}t)]\sigma(\lambda)\, d\lambda \right|.$$

In the interval $(\xi_l;\xi_{l+1}]$ we have

$$|\sigma(\lambda)| = |\sigma(\lambda)-\sigma(\xi_l)|$$
$$\leqq \max\{2\pi^{-1}[\lambda^{\frac12}-\xi_l^{\frac12}], [\rho(\lambda)-\rho(\xi_l)]\} \leqq 2\pi^{-1}[\xi_{l+1}^{\frac12}-\xi_l^{\frac12}],$$

hence by using $\xi_{l+1} < c\xi_l$ we get the evaluation

$$|u_l^{(k+3)}(t)| \leqq \tfrac{1}{2\pi} c^{\frac12 k}(t+k+1)\xi_l^{\frac12(k-1)}(\xi_{l+1}-\xi_l)^2.$$

Reference to (4.5) finishes the proof.

Remark 2. It is clear from the proof that the sum in (4.5) essentially bounds the $(k+3)$-th derivative of $a(t)$. Again, by using the relations between $q(t)$ and $a(t)$ we obtain the following statements: if the sum is finite for all k, $a(t)$ and $q(t)$ are of class C^∞; if the sums are bounded by $CM^k k!$, $a(t)$ and $q(t)$ are analytic.

Condition (4.5) will be the one to be used in our examples. It is of interest however to replace it by a condition more directly connected with the structure of the measure μ.

Let $(\lambda_i;\lambda'_i)$ be all the component intervals of $(1;\infty)-S$ where S is the closed support of μ.

THEOREM 3. *The existence of a sequence $\{\xi_l\}$ satisfying the requirements of Theorem 2 is equivalent to the condition*

$$\sum_i \lambda_i^{\frac12(k-1)}(\lambda'_i-\lambda_i)^2 < \infty. \tag{4.6}$$

Proof. Necessity of (4.6). Since the intervals $(\lambda_i;\lambda'_i)$ cannot contain any of the intervals $(\xi_l;\xi_{l+1}]$, each of the former must be contained in one of the latter or in the union of two consecutive intervals of the latter kind. If $\sum^{(l)}$ denotes summation over all i such that $(\lambda_i;\lambda'_i)\subset(\xi_l;\xi_{l+2}]$ we get for $l\geqq 1$ (constant c satisfying $\xi_{l+1}\leqq c\xi_l$)

$$\sum{}^{(l)} \lambda_i^{\frac{1}{2}(k-1)}(\lambda'_i - \lambda_i)^2 \leqq c^{k-1}\xi_l^{\frac{1}{2}(k-1)}(\xi_{l+2} - \xi_l)^2$$
$$\leqq 2c^{k-1}[\xi_l^{\frac{1}{2}(k-1)}(\xi_{l+1} - \xi_l)^2 + \xi_{l+1}{}^{\frac{1}{2}(k-1)}(\xi_{l+2} - \xi_{l+1})^2],$$
$$\sum{}^{(0)} \lambda_i^{\frac{1}{2}(k-1)}(\lambda'_i - \lambda_i)^2 \leqq \xi_2^{\frac{1}{2}(k+3)}.$$

By adding up these inequalities we obtain (4.6) from (4.5).

Sufficiency of (4.6). We choose a sequence $\{\eta_l\}$ satisfying all the conditions of Theorem 2 (when ξ_l is replaced by η_l) except perhaps the condition $\mu(\eta_{l+1}) - \mu(\eta_l) > 0$. This is obviously always possible.[8]

We arrange the sequence of intervals $(\eta_n; \eta_{n+1}]$ into successive groups of consecutive intervals so that the total measure μ of each group be positive. We define these groups successively, each time choosing the *smallest* group with positive total measure. All the groups are finite (because of (4.6)) and hence they form consecutive intervals $(\eta_{n_l}; \eta_{n_{l+1}}]$, $l = 0, 1, 2, \cdots$, with the following properties: 1°: n_l are strictly increasing integers, $n_0 = 0$, $\eta_{n_1} \geqq 1$; 2°: if $n_{l+1} - 1 > n_l$ and $l \geqq 1$, there exists an interval $(\lambda_{i_l}; \lambda'_{i_l})$ such that $\eta_{n_l - 1} < \lambda_{i_l} \leqq \eta_{n_l} < \eta_{n_{l+1}-1} \leqq \lambda'_{i_l} < \eta_{n_{l+1}}$. We now put $\xi_l = \eta_{n_l}$. If $n_{l+1} - 1 = n_l$ then the term $\xi_l^{\frac{1}{2}(k-1)}(\xi_{l+1} - \xi_l)^2$ is part of the sum (4.5) written for η_l and the sum of these terms is finite. If $n_{l+1} - 1 > n_l$, $l \geqq 1$, then by 2°: $\xi_l^{\frac{1}{2}(k-1)}(\xi_{l+1} - \xi_l)^2 \leqq \eta_{n_l}{}^{\frac{1}{2}(k-1)} 2[(\lambda'_{i_l} - \lambda_{i_l})^2 + (\eta_{n_{l+1}} - \eta_{n_{l+1}-1})^2]$. Since (4.6) implies that for some constant c_1, $\lambda'_i \leqq c_1\lambda_i$ it follows that

$$\xi_l^{\frac{1}{2}(k-1)}(\xi_{l+1} - \xi_l)^2 \leqq 2c_1{}^{\frac{1}{2}(k-1)}\lambda_{i_l}{}^{\frac{1}{2}(k-1)}(\lambda'_{i_l} - \lambda_{i_l})^2 + 2\eta_{n_{l+1}-1}{}^{\frac{1}{2}(k-1)}(\eta_{n_{l+1}} - \eta_{n_{l+1}-1})^2;[9]$$

hence the sum of these terms is again finite by (4.6) and by (4.5) written for η_l. Thus (4.5) is proved for $\{\xi_l\}$ and obviously this sequence satisfies all the other conditions of Theorem 2.

Remark 3. The preceding proof can be extended easily to prove that the sum (4.6) admits of a majoration $CM^kk!$ if and only if there exists a sequence ξ_l satisfying the conditions of Theorem 2 such that a similar majoration of the sum (4.5) is true.

5. We construct the first example announced in Section 1. Consider a sequence $\{\xi_l\}$ defined for all integers l with $-N < l < \infty$ and $0 \leqq N \leqq \infty$.

---

[8] For instance for a suitable $\tau > (\log 2)^{-1}$, we can put $\eta_l = \tau \log(l + 1)$. For this choice $\sum \eta_l{}^{\frac{1}{2}(k-1)}(\eta_{l+1} - \eta_l)^2$ is finite for all k and of order of magnitude $C_\epsilon \epsilon^k k!$ for all $\epsilon > 0$ and suitable constants C_ϵ.

[9] This evaluation is valid for $k \geqq 1$. For $k = 0$ it is obvious how it should be changed.

We suppose that $\xi_l < \xi_{l+1}$ for all l, $\xi_1 > 1$, and that $\xi_l \to \infty$ when $l \to \infty$, and finally that

$$(5.1) \qquad \sum_1^\infty \xi_l^{-\frac{1}{2}} (\xi_{l+1} - \xi_l)^2 < \infty.$$

In each interval $(\xi_l; \xi_{l+1})$ take a closed set S_l dense in itself, with $|S_l| = 0$. We define a Sturm-Liouville problem $-x'' + (q - \zeta)x = 0$ with boundary condition $x'(0) = \operatorname{tang} \alpha\, x(0)$, with continuous $q(t)$, and $-\frac{1}{2}\pi < \alpha < \frac{1}{2}\pi$, such that the spectral measure ρ_α is a pure singular continuous measure with support $S = \cup S_l$. This problem will be such that for any other boundary condition with $\beta \neq \alpha \bmod \pi$, ρ_β will be a pure point measure concentrated on isolated points in the complement of S. To this effect we first construct for each S_l a continuous singular measure μ_l with support S_l such that for each $\xi \in S_l$ the upper derivative of μ_l at ξ is $= \infty$. It is well known that such measures exist.

We then define $\mu = \sum \mu_l$. It is clear that for support, μ has the set $S = \cup S_l$ and that for each $\xi \in S$, $\bar{d}\mu(\xi)/d\lambda = \infty$. If we now construct, as in Theorem 2, the measure $\rho = \rho_\alpha$ and the corresponding Sturm-Liouville problem, the measure ρ_α will be pure continuous singular with support S and $\bar{d}\rho_\alpha(\xi)/d\lambda = \infty$ for each $\xi \in S$.

Consider now any $\beta \neq \alpha \bmod \pi$ and apply statement III of Section 2 to $m_\alpha(\zeta) + \operatorname{cotg} \gamma$, where $\gamma = \beta - \alpha$. We see that

$$\limsup_{\eta \searrow 0} \operatorname{Im}[m_\alpha(\xi + i\eta) + \operatorname{cotg} \gamma] = \infty$$

for every $\xi \in S$. It follows from formula (3.3) that

$$\liminf_{\eta \searrow 0} | m_\beta(\xi + i\eta) - \operatorname{cotg} \gamma | = 0$$

for $\xi \in S$. By definitions (2.4^0) and (2.4^1) supports $S^0{}_\beta$ and $S^1{}_\beta$ are disjoint from S. From Theorem 1 we know that $S^0{}_\beta = 0$ and hence $S^1{}_\beta$ is the whole support of the measure ρ_β and lies in the complement of S. But on this complement the function $m_\alpha(\zeta) + \operatorname{cotg} \gamma$ is analytic and real with simple zeros, one at most in each of the components of the complement of S. Hence $S^1{}_\beta$ is the set of all zeros of $m_\alpha(\zeta) + \operatorname{cotg} \gamma$ in the complement of S. This proves all our assertions.

Remark. If we choose the sequence ξ_l so that $\sum_1^\infty \xi_l^{\frac{1}{2}(k-1)} (\xi_{l+1} - \xi_l)^2$ be finite for some $k \geqq 1$, or finite for all $k \geqq 0$ and majorated by $CM^k k!$, we would get in our example $q(t)$ k-times continuously differentiable, or analytic, respectively.

6. Before we construct the second example we prove the following theorem

THEOREM 4. *Consider the Sturm-Liouville equation* (1.1) *and two boundary conditions corresponding to* $\alpha \neq \beta \operatorname{mod} \pi$, $\gamma = \beta - \alpha$. *In order that* ξ *be in the point spectrum relative to the boundary condition* (1.2_β) *it is necessary and sufficient that* $\int_{-\infty}^{\infty} (\lambda - \xi)^{-2} d\rho_\alpha(\lambda) < \infty$ *and that* $m_\alpha(\xi) + \operatorname{cotg}\gamma = 0$.

Proof. By statement IV in Section 2 we know that for the point ξ to have a positive mass $\rho_\beta\{\xi\}$ it is necessary and sufficient that

$$\lim (\xi - \zeta)[m_\beta(\zeta) - \operatorname{cotg}\gamma] = \rho_\beta\{\xi\}$$

when $\zeta \to \xi$ in any angle. By (3.3) this condition is equivalent to

$$\lim (\zeta - \xi)^{-1}[m_\alpha(\zeta) + \operatorname{cotg}\gamma] = [\sin^2 \gamma \rho_\beta\{\xi\}]^{-1} < \infty.$$

In the last equation put $\zeta = \xi + i\eta$, $\eta \searrow 0$. Using the expression (3.2) we obtain

$$\operatorname{Re}[(i\eta)^{-1}\{\operatorname{cotg}\gamma + \sigma_\alpha + \int_{-\infty}^{\infty} [(\lambda - \xi - i\eta)^{-1} - \lambda(\lambda^2 + 1)^{-1}] d\rho_\alpha(\lambda)\}]$$
$$= \int [(\lambda - \xi)^2 + \eta^2]^{-1} d\rho_\alpha(\lambda).$$

The last integrand is a decreasing function of η for each λ and it follows that for $\eta \searrow 0$ the integral converges to $\int (\lambda - \xi)^{-2} d\rho_\alpha(\lambda)$. Thus we get

$$\int (\lambda - \xi)^{-2} d\rho_\alpha(\lambda) = [\sin^2 \gamma \rho_\beta\{\xi\}]^{-1} < \infty.$$

It follows further that

$$\lim_{\eta \searrow 0} \{\operatorname{cotg}\gamma + \sigma_\alpha + \int [(\lambda - \xi - i\eta)^{-1} - \lambda(\lambda^2 + 1)^{-1}] d\rho_\alpha(\lambda)\}$$
$$= \operatorname{cotg}\gamma + \sigma_\alpha + \int [(\lambda - \xi)^{-1} - \lambda(\lambda^2 + 1)^{-1}] d\rho_\alpha(\lambda) = 0.$$

This gives the necessity of our conditions. If the conditions are satisfied, then by the last equation we can write

$$\operatorname{cotg}\gamma + m_\alpha(\zeta) = \int [(\lambda - \zeta)^{-1} - (\lambda - \xi)^{-1}] d\rho_\alpha(\lambda)$$
$$= (\zeta - \xi) \int [1/(\lambda - \zeta)(\lambda - \xi)] d\rho_\alpha(\lambda).$$

In view of our conditions, the quotient of this expression by $(\zeta - \xi)$ converges to $\int (\lambda - \xi)^{-2} d\rho_\alpha(\lambda)$ when $\zeta \to \xi$ in any angle. Then by reversing the previous argument we arrive at the conclusion that ρ_β has a positive mass at the point ξ.

In our next example we choose an arbitrary enumerable set S dense on the whole real axis and construct a Sturm-Liouville problem such that for one α, ρ_α is a pure point measure with support S, whereas for all $\beta \neq \alpha \bmod \pi$, ρ_β will be a pure singular continuous measure. To this effect we will assign to each $\xi \in S$ a positive measure $\mu\{\xi\}$ in such a way that the total measure in any finite interval be finite, and for each real ξ the quotient $\mu(I)/|I|^2 \to \infty$ for at least one sequence of intervals I containing ξ and converging to ξ. If this condition is satisfied then obviously for each ξ we have $\int (\lambda - \xi)^{-2} d\mu(\lambda) = \infty$ and this will be true also for the equivalent measure $\rho = \rho_\alpha$ defined in Theorem 2 (by virtue of the last assertion in this theorem). Therefore by Theorem 4, for all $\beta \neq \alpha \bmod \pi$ there cannot be any point spectrum. On the other hand, there cannot be any absolutely continuous spectrum since there is none for ρ_α (see Theorem 1). Hence ρ_β is a pure continuous singular measure.

There remains the construction of the measure μ. We define it in each of the intervals $(l;l+1]$. Since the construction is completely similar in each interval we shall do it only for $(0;1]$. We consider consecutive subdivisions of the interval in 2^k subintervals, $I_{k,i} = (2^{-k}(i-1); 2^{-k}i]$, $i = 1, 2, \cdots, 2^k$, $k = 1, 2, 3, \cdots$. To each subinterval we assign a point $\xi_{k,i} \in S \cap I_{k,i}$ in the following way. We consider the points of S as arranged in a simple sequence, $\eta_1, \eta_2, \cdots$. To each of the intervals $I_{1,i}$ we assign the first element of the sequence $\{\eta_n\}$ which lies in this interval. If the $\xi_{k,i}$ are already defined for all $k < k_0$, we define $\xi_{k_0,i}$ as the first element in $\{\eta_n\}$ which lies in $I_{k_0,i}$ and which was not assigned to any previous interval. It is clear that in this way we will assign points $\xi_{k,i}$ to all intervals $I_{k,i}$ and the set of all $\xi_{k,i}$ is exactly equal to $S \cap (0;1]$. To the points $\xi_{k,i}$ we assign the measure $\mu\{\xi_{k,i}\} = 3^{-k}$. The total measure for the interval $(0;1]$ is then $\sum_1^\infty (2/3)^k = 2$.

Consider now any point ξ with $0 < \xi \leqq 1$ and the sequence of intervals I_{k,i_k} containing ξ. We have $\mu(I_{k,i_k}) > \mu\{\xi_{k,i_k}\} = 3^{-k}$, $|I_{k,i_k}| = 2^{-k}$, hence $\mu(I_{k,i_k})/|I_{k,i_k}|^2 > (4/3)^k \to \infty$ which finishes the proof.

Remark. Since in the above example, there are no intervals with μ-

measure 0, the developments of Section 4 allow us to construct a corresponding Sturm-Liouville problem with an analytic function $q(t)$.

UNIVERSITY OF KANSAS.

REFERENCES.

[1] E. A. Coddington and N. Levinson, *Theory of ordinary differential equations*, McGraw-Hill, New York, 1955.

[2] I. M. Gelfand and B. M. Levitan, "Determination of a differential equation by its spectral function," *Izviestia Ak. N. SSSR S. M.*, vol. 15 (1951), pp. 309-360.

[3] P. Hartman, "On the essential spectra of symmetric operators in Hilbert space," *American Journal of Mathematics*, vol. 75 (1953), pp. 229-240.

[4] M. Krein, "On a method of effective solution of an inverse boundary problem," *Dokl. Ak. N. SSSR*, vol. 94 (1954), pp. 987-990.

[5] C. R. Putnam, "On the continuous spectra of singular boundary value problems," *Canadian Journal of Mathematics*, vol. 6 (1954), pp. 420-426.

[6] S. Saks, *Theory of the Integral*, Warsaw-Lwow, 1937.

[7] H. Weyl, "Über gewöhnliche Differentialgleichungen mit Singularitäten . . .," *Mathematische Annalen*, vol. 68 (1910), pp. 222-269.

CHARACTERIZATION OF POSITIVE REPRODUCING KERNELS. APPLICATIONS TO GREEN'S FUNCTIONS.*[1]

By N. Aronszajn and K. T. Smith.

1. **Introduction.** The object of this paper is to prove a theorem which characterizes the proper functional Hilbert spaces whose reproducing kernels are positive. Since certain Green's functions are among the most important examples of reproducing kernels, it is natural that the general theorem leads to results about the positiveness of Green's functions.

Not all Green's functions are reproducing kernels; those and only those are which correspond to positive definite differential problems of sufficiently high order. This restriction on the order is avoided by use of the notion of pseudo-reproducing kernels for general Hilbert spaces. Details of the underlying theory of pseudo-reproducing kernels are not given, except for those basic definitions and properties which are necessary to make clear the application of the abstract theorem about positiveness.

Roughly speaking[2] in the case of a self-adjoint elliptic problem, the Green's function exists if the problem is well posed, i.e., if there exists no solution to the homogeneous equation satisfying the homogeneous boundary conditions (in other words if 0 is not an eigenvalue of the corresponding eigenvalue problem).[3] A Green's function exists and is a reproducing or pseudo-reproducing kernel if and only if the problem is definite positive (i.e., if all the eigenvalues are positive). Our abstract theorem about reproducing kernels gives necessary and sufficient conditions in order that a Green's function corresponding to a *positive definite problem* be positive. We do not know if there exist positive Green's functions corresponding to problems which are not positive definite.

In the case of second order problems the question of which Green's functions are positive is answered completely by a proof of the equivalence of the following statements: (i) the problem is positive definite and (ii) the

* Received January 21, 1957.

[1] Paper written under contract with the Office of Naval Research Nonr 583(04).

[2] The additional restrictions under which the following statements can actually be proved are stated in Section 4.

[3] This excludes from our considerations the Neumann function for $\Delta u = 0$; this function actually changes sign in the domain.

Green's function exists and is positive. In the case of higher order problems there are examples showing that (ii) does not follow from (i) even in case of the square of the Laplacian and the boundary conditions $u = \partial u/\partial n = 0$.[4] However, we give examples of fourth order problems where it can be proved by the abstract theorem that the Green's function is positive, and where a simple direct proof does not seem to lie near at hand. The results are not restricted to problems about domains in Euclidean space. They are stated and proved for problems about relatively compact domains in oriented differentiable manifolds.

The question of positiveness of Green's functions is a very old one. For special cases like the Laplace equation with Dirichlet boundary condition, it goes back to the classical papers of C. Neumann and H. A. Schwarz where the positiveness of the Green's function is an important tool in investigations. Since then the positiveness of the Green's function was proved by many authors in many special cases of second order problems. Since for such problems with Dirichlet boundary condition the positiveness of the Green's function is closely related with the maximum principle, we should mention here the work of E. Hopf [12] who investigated this principle for the most general second order equations. It would seem that the first formal recognition of the relation between the positiveness of the Green's function and the positive character of the differential problem of second order with the Dirichlet boundary condition can be found in the thesis of Bleuler [13].

Our results on second order problems include as quite special cases some of the results of Bergman and Schiffer [6],[5] and can be proved, in a heuristic manner at least, by following the lines set forth by Bergman and Schiffer in the special cases. The difficulty in this approach lies in the lack of general information about the sets of points on which solutions to the differential equation are zero. Our method does not require information of this kind.

2. Reproducing kernels. The main theorem of the paper, Theorem 1 below, gives necessary and sufficient conditions in order that the reproducing

[4] The problem of positiveness of the corresponding Green's function in case of two variables is the famous problem of Hadamard stated in the first decade of this century and solved negatively less than ten years ago by Duffin [7] and Garabedian [8].

[5] These authors consider only the differential operator $-\Delta + p$, $p > 0$, and consider boundary conditions slightly more special than ours; they consider explicitly only the case of 2 variables. However, they treat two questions related to the question treated here; when a Green's function is positive, and when one Green's function is larger than another. The second question is not discussed in this paper, but will be discussed from a similar general point of view in another paper.

kernel $K(x,y)$ of a proper functional Hilbert space be non-negative.[6] It is proved in [2] that condition (a) in the theorem below is necessary and sufficient in order that $K(x,y)$ be real.[7]

THEOREM 1. *In order that the reproducing kernel $K(x,y)$ of the proper functional Hilbert space $\mathfrak{F}$ be non-negative it is necessary and sufficient that $\mathfrak{F}$ have the two properties*

(a) *If $u \in \mathfrak{F}$, then $\bar{u} \in \mathfrak{F}$ and $\|\bar{u}\| = \|u\|$.*

(b) *For each real-valued $u \in \mathfrak{F}$ there exists $\tilde{u} \in \mathfrak{F}$ such that*

$$\tilde{u}(x) \geqq |u(x)| \text{ for all } x \text{ and } \|\tilde{u}\| \leqq \|u\|.$$

Proof. Assuming that $\mathfrak{F}$ has the properties (a) and (b), let $u \in \mathfrak{F}$, u be real-valued and let $\tilde{u}$ be such that (b) holds, and set $u^+ = \frac{1}{2}(\tilde{u} + u)$ and $u^- = \frac{1}{2}(\tilde{u} - u)$. Clearly u^+ and u^- are both non-negative, and $\tilde{u} = u^+ + u^-$ and $u = u^+ - u^-$. By using (a) it is proved easily that the scalar product of real-valued functions is real. Therefore, from $\|\tilde{u}\| \leqq \|u\|$ it follows that $(u^-, u^+) \leqq 0$, and hence that $(u^-, u) = (u^-, u^+) - (u^-, u^-) \leqq -\|u^-\|^2$. By virtue of the result from [2] quoted above, $K_y(x) = K(x,y)$ is real-valued for every y so that this inequality can be applied to $u = K_y$ to give $0 \leqq K_y^-(y) = (K_y^-, K_y) \leqq -\|K_y^-\|^2$. This is possible only if $K_y^- = 0$ so that $K_y = K_y^+ \geqq 0$.

Assuming that $K(x,y) \geqq 0$, for each real-valued $u \in \mathfrak{F}$, let u' and u'' be the projections of u and of $-u$ on the closed convex cone with vertex 0 generated by the set $\{K_y\}$. Then for every $\rho \geqq 0$ and every v in the cone, $\|u - u'\|^2 \leqq \|u - u' - \rho v\|^2$, from which it follows that for every v in the cone $(u' - u, v) \geqq 0$. Applying this to $v = K_y$, we obtain $u'(y) \geqq u(y)$ for every y. Similarly, $u''(y) \geqq -u(y)$ for every y. Since $K(x,y)$ is non-negative, both u' and u'' are non-negative, so that if $\tilde{u} = u' + u''$, then $\tilde{u}(x) \geqq |u(x)|$ for all x. All that remains is to show that $\|\tilde{u}\| \leqq \|u\|$. This was done in another paper [5], but since the proof is very short it bears repetition.

---

[6] A proper functional Hilbert space is a Hilbert space whose elements are functions on a basic set $\mathcal{E}$, such that the value of a function at any point in $\mathcal{E}$ is a continuous linear functional. To each proper functional Hilbert space $\mathfrak{F}$ there corresponds a function $K(x,y)$, called the reproducing kernel of $\mathfrak{F}$, defined on $\mathcal{E} \times \mathcal{E}$ and having the properties:

(a) for each $y \in \mathcal{E}$, the function $K_y(x) = K(x,y)$ belongs to $\mathcal{E}$.

(b) the reproducing property: for each $u \in \mathfrak{F}$, $u(y) = (u, K_y)$.

[7] If u is a complex valued function, $\bar{u}$ is its complex conjugate; i.e., $\bar{u}(x)$ is the complex conjugate of $u(x)$.

If ϕ', α, and ϕ'' are the angles between u and u', u' and u'', and u'' and $-u$, respectively, then, as $u-u'$ is orthogonal to u' and $-u-u''$ is orthogonal to u'', the inequality to be proved takes the form

$$\| u \|^2(\cos^2\phi' + \cos^2\phi'' + 2\cos\phi'\cos\phi''\cos\alpha) \leqq \| u \|^2,$$

which is easy to establish with the aid of the inequalities $0 \leqq \phi' \leqq \frac{\pi}{2}$, $0 \leqq \phi'' \leqq \frac{\pi}{2}$, and $\pi \leqq \phi' + \alpha + \phi''$.

Remark 1. If condition (a) holds for a set of functions which is dense in $\mathfrak{F}$, then condition (a) holds for all function in $\mathfrak{F}$. If condition (b) holds for a set of functions which is dense in the set of all real-valued functions in $\mathfrak{F}$, then condition (b) holds for all real-valued functions in $\mathfrak{F}$.

Remark 2. The following statement about matrices is a consequence of Theorem 1. *Let $\{a_{ij}\}$ be a real positive definite matrix. In order that every element in $\{a_{ij}\}^{-1}$ be non-negative it is necessary and sufficient that for each real vector u there exist $\tilde{u}$ such that $\tilde{u}_i \geqq |u_i|$ for all i and $\sum a_{ij}\tilde{u}_i\tilde{u}_j \leqq \sum a_{ij}u_iu_j$.*

3. Pseudo-reproducing kernels. The functional spaces which are of use in differential problems are proper functional Hilbert spaces in problems of order larger than the dimension of the underlying Euclidean space. In problems of smaller order the spaces are more general functional Hilbert spaces in which the functions are not defined everywhere.

A functional space is a normed linear class $\mathfrak{F}$ of functions on a basic set $\mathcal{E}$, each defined except on some exceptional set belonging to a hereditary σ-ring $\mathfrak{A}$ (the exceptional class of sets). It is assumed that the norm in $\mathfrak{F}$ has the property that if $\| u_n - u \| \to 0$, then, for some subsequence $\{u_{n_k}\}$, $u_{n_k}(x) \to u(x)$ except on a set in $\mathfrak{A}$. A complete discussion of functional spaces and functional completion can be found in [5].

Let μ be a σ-finite complete measure on a set $\mathcal{E}$. The class of measurable subsets of $\mathcal{E}$ of finite measure will be denoted by $\mathcal{E}'$, the sets in $\mathcal{E}'$ by x', y', etc. A *μ-measurable functional Hilbert space* is a functional Hilbert space $\mathfrak{F}$ on the basic set $\mathcal{E}$ such that: (a) all the exceptional sets are of measure 0, (b) each $u \in \mathfrak{F}$ is integrable on every $x' \in \mathcal{E}'$, (c) if $\| u \| \neq 0$, *then* $\int_{x'} u\, d\mu \neq 0$ for some $x' \in \mathcal{E}'$, and (d) for every $x' \in \mathcal{E}'$, $\int_{x'} u\, d\mu$ is a continuous linear functional of u.[8]

A μ-measurable functional Hilbert space $\mathfrak{F}$ determines a proper functional Hilbert space $\mathfrak{F}'$ whose basic set is the class $\mathcal{E}'$. $\mathfrak{F}'$ consists of all

[8] The definition is valid for general incomplete spaces. In case of a complete space, in particular in case of a Hilbert space, (d) follows from (a), (b), (c).

functions u' to which there corresponds some $u \in \mathfrak{F}$ such that $u'(x') = \int_{x'} u\, d\mu$ for every $x' \in \mathcal{E}'$. The mapping $u \to u'$ is one to one, and when (u', v') is defined to be (u, v), it becomes a Hilbert space isomorphism between $\mathfrak{F}$ and $\mathfrak{F}'$. The reproducing kernel for $\mathfrak{F}'$ is the function $K'(x', y') = (u_{y'}, u_{x'})$, where $u_{x'} \in \mathfrak{F}$ is such that for every $v \in \mathfrak{F}$, $\int_{x'} v\, d\mu = (v, u_{x'})$.

A pseudo-reproducing kernel for a μ-measurable functional Hilbert space $\mathfrak{F}$ is a function $K(x, y)$ with the property that for every x' and y' in $\mathcal{E}'$, $K(x, y)$ is integrable over $x' \times y'$, and $K'(x', y') = \int_{x'} \int_{y'} K(x, y) d\mu(y) d\mu(x)$. It is clear that this condition determines $K(x, y)$ a.e. on $\mathcal{E} \times \mathcal{E}$.[9]

Since $K'(x', y')$ is an additive function of the rectangle $x' \times y'$, it can be extended to an additive set function on the ring generated by the rectangles, so that by the Radon-Nikodym theorem, a necessary and sufficient condition in order that the μ-measurable functional Hilbert space $\mathfrak{F}$ have a pseudo-reproducing kernel is that the function $K'(x', y')$ be an absolutely continuous function of the rectangle $x' \times y'$.

If a measure μ and a function $K(x, y)$ integrable over all $x' \times y'$ are given, then a necessary and sufficient condition in order that $K(x, y)$ be a pseudo-reproducing kernel for some μ-measurable functional Hilbert space $\mathfrak{F}$ is that the function $K'(x', y') = \int_{x'} \int_{y'} K(x, y) d\mu(y) d\mu(x)$ be a positive matrix on $\mathcal{E}'$. If $K'(x', y')$ is a positive matrix, $\mathfrak{F}$ is constructed as follows. Let $\mathfrak{F}_0$ be the class of all linear combinations $u = \sum \alpha_i u_{y'_i}$, where $u_{y'_i}(x) = \int_{y'_i} K(x, y) d\mu(y)$, with the norm

$$\| u \|^2 = \sum \alpha_i \bar{\alpha}_j K'(y'_j, y'_i). \tag{3.1}$$

It can be proved that $\mathfrak{F}_0$ has a functional completion $\mathfrak{F}$, which is a μ-measurable functional Hilbert space.[10] Obviously, $K(x, y)$ is a pseudo-reproducing kernel for $\mathfrak{F}$.

The proper generalization of Theorem 1 to μ-measurable functional Hilbert spaces and pseudo-reproducing kernels is clear.

---

[9] The beginnings of the theory of measurable spaces and pseudo-reproducing kernels are in the notes [1]. These notions can be defined in an absolute way, independent of a measure given beforehand. The relative notions defined here are sufficient for our needs in the present paper and avoid some of the difficulties inherent in the absolute notions.

[10] The proof is not obvious. It will be given in the development of the general theory of measurable spaces and pseudo-reproducing kernels.

If we put $f = \sum \alpha_i \chi_{y'_i}$, where $\chi_{y'_i}$ is the characteristic function of y'_i, then (3.1) reads

(3.2) *If* $u = Kf = \int K(x,y)f(y)d\mu(y)$, *then* $\| u \|^2 = \int Kf(x)\overline{f(x)}d\mu(x)$.

Thus $K(x,y)$ is a pseudo-reproducing kernel if and only if it is the kernel of an integral operator which is positive in the most general sense, i.e., on the class of linear combinations of characteristic functions of measurable sets of finite measure.[11] The corresponding μ-measurable functional Hilbert space is the completion of the range of the integral operator in the norm (3.2). Theorem 1 has the following corollary about positive integral operators.

COROLLARY 1. *If K is a positive integral operator defined on the class of linear combinations of characteristic functions of sets of finite measure, then the kernel of K is non-negative almost everywhere if and only if it is real and for each real u in the range of K there exists $\tilde{u}$ in the completion of the range such that $\tilde{u}(x) \geqq |u(x)|$ almost everywhere and $\|\tilde{u}\| \leqq \|u\|$. The norm in question is that defined in* (3.2).

4. Green's functions. The kernels in which we are chiefly interested are the Green's functions for differential systems. Let A be a linear elliptic self-adjoint differential operator defined in a bounded domain D, and let $\{B_i\}$ be a normal system of boundary operators such that the system $(A;\{B_i\})$ is self-adjoint.[12] The differential system $(A;\{B_i\})$ defines a symmetric operator, which also will be called A, on the Hilbert space L^2 of functions which are square integrable over D. The domain of the operator A is the set of all sufficiently regular functions u on $\bar{D}$ which satisfy the boundary conditions $B_i u = 0$. If the closure of the operator A has a bounded inverse defined everywhere on L^2, and if the inverse is an integral operator, then the kernel $G(x,y)$ of this integral operator is called the Green's function of the system $(A;\{B_i\})$. It is clear that if $u = Gf$, then $\int_D Gf(x)\overline{f(x)}dx = \int_D u(x)\overline{Au(x)}dx$, so that the quadratic form in (3.2) is non-negative if and only if[13]

$$\| u \|^2 = \int_D Au(x)\overline{u(x)}\,dx, \tag{4.1}$$

[11] We treat the complex case. In the real case it must be assumed that $K(x,y) = K(y,x)$.

[12] We use the terminology introduced in Aronszajn and Milgram [4].

[13] In order to reduce the number of notations we write A instead of the closure of A in some of the formulas.

is non-negative for all u in the domain of A. Since it is assumed that the closure of A has a bounded inverse, if (4.1) is non-negative, there exists a constant $c > 0$ such that

$$\| u \|^2 \geqq c \int_D | u |^2 dx \text{ for } u \text{ in the domain of } A;$$

that is, the differential system $(A; \{B_i\})$ is *positive definite*.

In classical usage the term Green's function is restricted to kernels $G(x,y)$ which are sufficiently regular: if the order of the operator A is $2m$, if its coefficients are sufficiently differentiable, and if the boundary ∂D is sufficiently smooth, then, for fixed $y \in \bar{D}$, $G(x,y)$, as a function of x, should be at least of class C^{2m} in $\bar{D} - \{y\}$. Recent advances in the theory of elliptic partial differential equations have made it possible to prove, in a wide variety of cases, not only the existence of the Green's function as we have defined it, but also the regularity required in the classical definition. The proof, especially the proof of regularity, is based on the property of coerciveness, which has to do with the relation between the quadratic form $\int_D Au\,\bar{u}\,dx$ and the standard m-norm, defined as follows

$$\| u \|_m{}^2 = \sum_{k=0}^{m} \sum_{k_1 + \cdots + k_n = k} k!(k_1! \cdots k_n!)^{-1} \int_D | \partial^k u / \partial x_1^{k_1} \cdots \partial x_n^{k_n} |^2 \, dx. \tag{4.2}$$

The quadratic form $\int_D Au\,\bar{u}\,dx$ is said to be *coercive* on the domain of A if there exist constants $c' \geqq 0$ and $N > 0$ such that

$$\int_D Au\,\bar{u}\,dx + c' \| u \|_0{}^2 \geqq N \| u \|_m{}^2 \text{ for all } u \text{ in the domain of } A.\text{[14]} \tag{4.3}$$

If the quadratic form $\int_D Au\,\bar{u}\,dx$ is coercice on the domain of A, the following statement can be proved.

If the coefficients of A and the boundary of D are of class C^{2m}, then the existence of the Green's function $G(x,y)$ in the sense given above is equivalent to the fact that 0 is not an eigenvalue of the closure of A. Furthermore, if $G(x,y)$ exists, the integral operator with kernel $G(x,y)$ is completely continuous in the space L^2.[15]

[14] The first coerciveness inequality was proved by Gårding [9] for the case of Dirichlet boundary conditions. The general notion of coerciveness was introduced and investigated by Aronszajn [3]. A more complete presentation of the subject will be given by the same author in a forthcoming paper.

[15] Actually, the above statement is true even when the coefficients and the boundary are only $C^{m-1,1}$; however it is necessary then to consider boundary conditions of order $\geqq m$ in a suitable generalized sense. If we assume the coefficients of A to be of suffi-

If the Green's function of the system $(A;\{B_i\})$ exists and if the system is positive definite, then we see from the last section that the Green's function is a pseudo-reproducing kernel. The corresponding functional Hilbert space is the completion of the domain of A in the norm (4.1). If the norm (4.1) is coercive, it can be proved that the domain of A has a perfect functional completion, the functions of which have derivatives a.e. in the ordinary sense of orders $\leqq m$ in D and of orders $\leqq m-1$ on the boundary of D. This completion contains all functions of class $C^{(m-1,1)}$ (i.e. functions of class C^{m-1} with Lipschitz $(m-1)$-st derivatives) which satisfy the stable boundary conditions, that is those of order $\leqq m-1$. Therefore, Theorem 1 or Corollary 1 give a necessary and sufficient condition in order that the Green's function be non-negative.

It is clear that the foregoing considerations are valid for differential systems considered in a relatively compact subdomain of a differentiable manifold.

5. Green's functions for differential problems of order 2. We consider a real elliptic linear differential operator A of order 2 on an oriented differentiable manifold M^n of class C^3. Such an operator (or its negative) is expressible in each coordinate patch in the form

$$Au = -\sum a^{ij}\partial^2 u/\partial x^i \partial x^j + \textit{terms of lower order}$$

where a^{ij} is a real symmetric contravariant tensor of rank 2 which forms a positive definite matrix at every point. It is assumed that the tensor a^{ij} is of class C^2, that the coefficients of the first derivatives of u (which do not form a tensor) are of class C^1, and that the coefficient of u is continuous. It is assumed further that the operator A is self-adjoint with respect to some positive density $\rho(x)dx^1\cdots dx^n$ of class C^2.[16]

If $\{a_{ij}\}$ denotes the matrix inverse to $\{a^{ij}\}$ and if a denotes the determinant of $\{a_{ij}\}$, then $g_{ij}=a^{-1/n}\rho^{2/n}a_{ij}$ is a Riemannian metric on M^n. Henceforth M^n is considered as a Riemannian manifold with this metric.[17] If g

ciently high Class C^N, it can be proved that $G(x,y)$ exists in the classical sense. A proof of this fact will be published elsewhere. For special types of boundary conditions, including the Dirichlet boundary conditions, a proof of the regularity was given by Nirenberg [10]. A sketchy proof of the statement in the text was given in [11].

[16] This means that $\rho(x)dx^1\cdots dx^n$ is an exterior differential form of rank n, that in each coordinate patch $\rho(x)$ is positive and of class C^2, and that for every pair of functions u and v of class C^2 and 0 outside a compact set,

$$\int Au\bar{v}\rho(x)dx^1\cdots dx^n=\int u\overline{Av}\rho(x)dx^1\cdots dx^n.$$

[17] We have assumed that the coefficients of A are somewhat more regular than is strictly necessary, in order to be able to use this metric, thereby simplifying a number of the formulas considerably.

denotes the determinant of g_{ij} then $g^{\frac{1}{2}} = \rho$, so A is expressible in any coordinate patch in the form

$$Au = -\sum g^{-\frac{1}{2}} \partial (g^{\frac{1}{2}} a^{ij} \partial u / dx^i) / \partial x^j + \alpha u$$

where α is a continuous real valued function.

Let D be a relatively compact domain in M^n, and let M^{n-1} be a portion of the boundary, ∂D, which is a submanifold of M^n of class C^2. It can be proved that the most general normal system of boundary operators at M^{n-1} which is self-adjoint relative to A is a system composed of a single operator, either the operator $B_0 u = u$ or the operator $B_1 u = \partial u / \partial \nu + bu$, where b is an arbitrary continuous real valued function on M^{n-1}, and where $\partial / d\nu$ is the interior normal derivative in the metric g_{ij}.[18]

It is assumed that the boundary ∂D is a submanifold of M^n of class $C^{(0,1)}$, and that ∂D is also the boundary of $M^n - \bar{D}$. It is assumed in addition that ∂D is piecewise of class C^2 in the following sense: there exist a finite number of disjoint submanifolds M_i^{n-1} of M^n such that $\partial D = \cup \bar{M}_i^{n-1}$; for each i, $\bar{M}_i^{n-1}$ is contained in a submanifold N_i^{n-1} of M^n of class C^2; and in the usual measure on N_i^{n-1}, $\bar{M}_i^{n-1} - M_i^{n-1}$ has measure 0.

Let B be a boundary operator at ∂D obtained by choosing for each M_i^{n-1} either the operator B_0 or one of the operators B_1, (with b continuous on $\bar{M}_i^{n-1}$) and consider the bilinear form

$$(5.1) \quad Q(u,v) = \int_D a^{ij} \partial u / \partial x^i \, \partial \bar{v} / \partial x^j \, g^{\frac{1}{2}} \, dx + \int_D \alpha u \bar{v} g^{\frac{1}{2}} \, dx - \int_{\partial D} p(s) k u \bar{v} \, ds$$

where k is the function on ∂D which on M_i^{n-1} is equal to 0 if the boundary operator chosen on M_i^{n-1} is B_0 and is equal to $-b$ if the boundary operator chosen on M_i^{n-1} is B_1, and where $p(s)$ is the value of the characteristic polynomial of A for the unit normal to ∂D at the point s. If u and v are functions of class C^2 in a neighborhood of $\bar{D}$ and if $Bu = Bv = 0$, then

$$(5.2) \qquad Q(u,v) = \int_D A u \bar{v} g^{\frac{1}{2}} \, dx.$$

The quadratic form $Q(u,u)$ is easily proved to be coercive, even on the class of all functions, a fortiori on the domain of A. Therefore, if the differential system $(A;B)$ is positive definite, then the Green's function exists

[18] The terminology used here is that introduced in Aronszajn and Milgram [4]; and all of the calculations used in finding the form of self-adjoint systems are based on the general theory of differential operators on Riemannian manifolds developed there. The operator B_1 has a particularly simple form because of the special choice of the Riemannian metric on M^n.

and is a pseudo-reproducing kernel for the completion of the domain of A with the norm $Q(u,u)^{\frac{1}{2}}$. In view of the statements in the last section about this functional space, the scalar product in the complete space is still given by (5.1). Furthermore, the class of Lipschitz functions satisfying the stable boundary conditions (i.e. $B_0 u = 0$ whenever $B = B_0$) forms a dense subspace of the completion. For a function u of this class, $|u|$ is clearly also of this class, and can be taken as the function $\tilde{u}$ in condition (b) of Theorem 1, since $Q(|u|,|u|) = Q(u,u)$. This leads to

THEOREM 2. *The three statements below are equivalent.*

(i) *The system $(A;B)$ is positive definite.*

(ii) *The Green's function for the system $(A;B)$ exists and is a pseudo-reproducing kernel.*

(iii) *The Green's function for the system $(A;B)$ exists and is non-negative.*

Proof. The equivalence of statements (i) and (ii) and the fact that they imply statement (iii) are already proved. We assume therefore, that the system $(A;B)$ is not positive definite, but that the Green's function $G(x,y)$ exists, and we show that $G(x,y)$ cannot be non-negative.

According to the part of the theorem which is proved, the Green's function $G_\rho(x,y)$ for the system $(A+\rho;B)$ exists and is non-negative, provided the number $\rho > 0$ is sufficiently large. G_ρ defines a completely continuous integral operator on L^2. Let $1/\mu$ be its largest eigenvalue. By a well-known theorem of Jentsch, $1/\mu$ is positive and there exists a non-negative corresponding eigenfunction f. It is easily verified that $\mu - \rho < 0$ and that $f = (\mu - \rho) Gf$. Obviously, therefore, G cannot be non-negative.

6. Green's functions for differential problems of Order 4. The examples of Duffin and Garabedian mentioned in the introduction show that there is no condition resembling the conditions (i) and (ii) of Theorem 2 which always will ensure that the Green's function of a problem of order 4 is positive. In this section we do not attempt to treat the general question; we only present some interesting special cases.

We assume that M^n is an oriented differentiable manifold of sufficiently high class, that D is a domain in M^n with a sufficiently regular boundary, and that A is a second order differential operator of the type considered in the last section with sufficiently regular coefficients. M^n is given a Riemannian metric as before. We consider differential system of the form $(A^2;B_0,B_2)$ and $(A^2;B_1,B_3)$, where

$$B_0 u = u,$$
$$B_1 u = \partial u/\partial \nu + bu,$$
$$B_2 u = -cp(s)^{-2}\,\partial u/\partial \nu + p(s)^{-1} Au,$$
$$B_3 u = dp(s)^{-2} u + p(s)^{-1} B_1 A u,$$

in which b c, and d are sufficiently regular functions on ∂D. Letting A_0 be the operator on L^2 corresponding to the first system, and A_1 be the operator corresponding to the second, we define

$$Q_0(u,u) = \int_D |Au|^2 g^{\frac{1}{2}}\,dx - \int_{\partial D} c\,|\partial u/\partial \nu|^2\,ds,$$
$$Q_1(u,u) = \int_D |Au|^2 g^{\frac{1}{2}}\,dx - \int_{\partial D} d\,|u|^2\,ds.$$

If u belongs to the domain of A_0, then $\int_D A^2 u\bar{u}\,dx = Q_0(u,u)$ and if u belongs to the domain of A_1, then $\int_D A^2 u\bar{u}\,dx = Q_1(u,u)$. Furthermore, it is not difficult to prove that $Q_0(u,u)$ is coercive on the domain of A_0 and $Q_1(u,u)$ is coercive on the domain of A_1. We get the following

THEOREM 3. *If both systems $(A;B_0)$ and $(A^2;B_0,B_2)$ are positive definite and if $c \geqq 0$, then the Green's function for $(A^2;B_0,B_2)$ exists and is non-negative. If both systems $(A;B_1)$ and $(A^2;B_1,B_3)$ are positive definite and if $d \geqq 0$, then the Green's function for $(A^2;B_1,B_3)$ exists and is non-negative.*

Proof. Only the first statement will be proved; the second is proved similarly.

The domain of A_0 with the norm $Q_0(u,u)^{\frac{1}{2}}$ has a functional completion whose elements are the functions of the form $u(x) = G_0 f(x)$, where G_0 is the Green's function for the system $(A;B_0)$ and where f is square integrable. Clearly

$$Q_0(u,u) = \int_D |f|^2 g^{\frac{1}{2}}\,dx - \int_{\partial D} c\,|\partial u/\partial \nu|^2\,ds.$$

Since $G_0 \geqq 0$, $\partial G_0/\partial \nu \geqq 0$. Hence the function $\tilde{u}(x) = G_0|f|(x)$ has the properties $\tilde{u}(x) \geqq |u(x)|$ and $\partial\tilde{u}/\partial\nu \geqq |\partial u/\partial \nu|$. Since also $c \geqq 0$ it follows that $Q_0(\tilde{u},\tilde{u}) \leqq Q_0(u,u)$.

Remark. If $(A;B_0)$ is positive definite, then, for sufficiently small c, $(A^2;B_0,B_2)$ is also positive definite. If $c = 0$, the statement of the last theorem is trivial, since Green's function for the system $(A^2;B_0,1/p(s)A)$ is

$\int_D G_0(x,z)G_0(z,y)g^{\frac{1}{2}}dz$. A similar remark holds for the system $(A^2;B_1,B_3)$ and the function d.

UNIVERSITY OF KANSAS.

REFERENCES.

[1] N. Aronszajn, "Les noyaux pseudo-reproduisants. Noyaux pseudo-reproduisants et complétion des classes hilbertiennes." "Complétion fonctionnelle de certaines classes hilbertiennes." "Propriétés de certaines classes hilbertiennes completées," *Comptes Rendus*, vol. 226 (1948), pp. 456, 537, 617, 700.

[2] ———, "Theory of reproducing kernels," *Transactions of the American Mathematical Society*, vol. 68 (1950), pp. 337-404.

[3] ———, *On coercive integro-differential quadratic forms*, Conference on Partial Differential Equations, University of Kansas, Summer, 1954.

[4] N. Aronszajn and A. N. Milgram, "Differential operators on Riemannian manifolds," *Rendiconti del Circolo Matematico di Palermo*, ser. 2, vol. 2 (1953), pp. 1-61.

[5] N. Aronszajn and K. T. Smith, "Functional spaces and functional completion," *Annales de l'Inst. Fourier de Grenoble*, vol. 6 (1955-56), pp. 125-185.

[6] S. Bergman and M. Schiffer, "Kernel functions in the theory of partial differential equations of elliptic type," *Duke Mathematical Journal*, vol. 15 (1948), pp. 535-566.

[7] R. J. Duffin, "On a question of Hadamard concerning superbiharmonic functions," *Journal of Mathematics and Physics*, vol. 27 (1949), pp. 253-258.

[8] P. R. Garabedian, "A partial differential equation arising in conformal mapping," *Pacific Journal of Mathematics*, vol. 1 (1951), pp. 485-524.

[9] L. Gârding, "Le problème de Dirichlet pour les équations aux derivées partielles élliptiques dans des domaines bornés," *Comptes Rendus*, Paris, vol. 233 (1951), pp. 1554-1556.

[10] L. Nirenberg, "Remarks on strongly elliptic partial differential equations," *Communications on Pure and Applied Mathematics*, vol. 8 (1955), pp. 649-675.

[11] K. T. Smith, *Functional spaces, functional completion, and differential problems*, Conference on Partial Differential Equations, University of Kansas, Summer, 1954.

[12] E. Hopf, "Elementare Bemerkungen über die Lösungen partieller Differentialgleichungen zweiter Ordnung vom elliptischen Typus," *Sitzungsberichte der Preussischen Akademie der Wissenschaften*, vol. 1927, pp. 147-152.

[13] K. Bleuler, "*Ueber den Rolle'schen Satz für den Operator $\Delta u + \lambda u$ und die damit zusammenhängenden Eigenschaften der Green'schen Funktion*," Zürich (Dissertation), 1941-1942.

GROUPS WHICH ACT ON S^n WITHOUT FIXED POINTS.*

By JOHN MILNOR.[1]

We will study the question of which finite groups can act on an n-sphere without fixed points. The main result (Section 1) is that any element of order two in such a group must lie in the center. A general survey of the problem is given in Section 2. The special case $n = 3$ is considered in Section 3.

An equivalent formulation of the problem is the following. Which groups can occur as the fundamental groups of manifolds M^n for which the universal covering space $\tilde{M}^n$ is homeomorphic to S^n. More generally one might only require that $\tilde{M}^n$ have the homotopy type of S^n. For example, any 3-manifold with a finite fundamental group satisfies this more general condition. (Such manifolds are classified up to homotopy type by the fundamental group together with the k-invariant (see Olum [3] Theorem IV.))

The author is indebted to helpful discussions with A. Shapiro.

1. Elements of order 2. A known theorem asserts that, for every map $f: S^n \to S^n$ of odd degree, there exists a pair of antipodal points whose images under f are also antipodal. (*Proof.* Let T denote the antipodal map, and suppose that $f(x)$, $f(Tx)$ are never antipodal. Let $g(x)$ denote the midpoint of the shortest great circle arc joining $f(x)$ to $f(Tx)$. Since $g(x) = g(Tx)$, the map $g: S^n \to S^n$ can be factored through n-dimensional projective space, and therefore has even degree. But since g is homotopic to f, this contradicts the assumption that f has odd degree.)

We will need the following generalization, which will be proved by a similar method.

THEOREM 1. *Let M^n be a manifold having the mod two homology of the n-sphere, and let $T: M^n \to M^n$ be a map of period two without fixed points. Then for every map $f: M^n \to M^n$ of odd degree there exists a point $x \in M^n$ such that $Tf(x) = fT(x)$.*

---

* Received April 2, 1957.

[1] The author holds an Alfred P. Sloan fellowship.

As an immediate consequence we have:

Corollary 1. *Suppose that the group G acts without fixed points on a manifold M^n having the mod two homology of the n-sphere. Then any element of order two in G belongs to the center.*

Proof. Let $T \in G$ have order two, and let U be any other element of G. Then Theorem 1 implies that we can find a point $x \in M^n$ with $TU(x) = UT(x)$. Since G acts without fixed points, this implies that $TU = UT$, which proves the corollary.

As an example we have

Corollary 2. *The symmetric group on three elements cannot operate on the n-sphere without fixed points.*

Remark. It is known that the group $Z_2 + Z_2$ cannot operate on S^n without fixed points. (See Section 2.) Therefore Corollary 1 implies that the group G can have at most one element of order 2. Conversely if G has only one element of order 2, then this element clearly belongs to the center.

Proof of Theorem 1. Let A denote the set of all points (x, Tx) in $M^n \times M^n$. Let $M^n * M^n$ denote the symmetric product, and let A' denote the set of points $\{x, Tx\}$ in $M^n * M^n$. If $f\colon M^n \to M^n$ satisfies $Tf(x) \neq fTx$ for all x, then there is a commutative diagram

$$\begin{tikzcd}
 & M_n \\
M^n \arrow[ur, "f"] \arrow[r, "f_1"] \arrow[d, "i_1"] & M^n \times M^n - A \arrow[u, "p_1"'] \arrow[d, "i_2"] \\
M^n/T \arrow[r, "f_2"] & M^n * M^n - A'
\end{tikzcd}$$

where i_1 and i_2 are the natural identification maps, and where

$$f_1(x) = (f(x), f(Tx)), \qquad f_2\{x, Tx\} = \{f(x), f(Tx)\}, \qquad p_1(x, y) = x.$$

Thus we are almost able to factor f through the "projective space" M^n/T, except that the map i_2 goes in the wrong direction.

LEMMA 1. *The homomorphism*

$$i_{2*}: H_k(M^n \times M^n - A) \to H_k(M^n * M^n - A')$$

is an isomorphism for all dimensions k.

(Singular homology with the coefficient group Z_2 is to be understood.)

Proof. It will first be shown that the projection $p_1: M^n \times M^n - A \to M^n$ is a locally trivial fibre map with fibre $p_1^{-1}(x) \approx M^n - Tx$. Given any closed n-cell $\bar{C}$ in M^n with interior C, choose a homeomorphism g of $T\bar{C}$ onto the unit ball $\bar{B}$ in euclidean n-space. Define the map $\lambda: B \times \bar{B} \to \bar{B}$ by

$$\lambda(\vec{v}, \vec{w}) = \vec{w} + (1 - \| \vec{w} \|)\vec{v},$$

where B denotes the interior of $\bar{B}$. Note that for each fixed vector $\vec{v} \in B$ the correspondence $\vec{w} \to \lambda(\vec{v}, \vec{w})$ defines a homeomorphism of $\bar{B}$ onto itself leaving the boundary fixed and carrying the center $\vec{0}$ into $\vec{v}$. The required local product structure

$$\phi: C \times (M^n - g^{-1}(\vec{0})) \to p_1^{-1}(C)$$

is now defined by

$$\phi(x, y) = (x, y) \text{ if } y \notin TC, \quad \phi(x, y) = (x, g^{-1}\lambda(gTx, gy)) \text{ if } y \in T\bar{C}.$$

It is not hard to show that ϕ is a homeomorphism.

Thus p_1 is a fibre map. According to the Lefschetz duality theorem, the fibre $M^n - Tx$ has the homology of a point. Therefore the induced homomorphism

$$p_{1*}: H_*(M^n \times M^n - A) \to H_*(M^n)$$

is an isomorphism. Let Δ denote the diagonal in $M^n \times M^n - A$ and let Δ' denote its image in $M^n * M^n - A'$. Since the diagonal map $M^n \to \Delta \subset M^n \times M^n - A$ is a cross-section of the fibre map p_1, it follows that the group $H_*(M^n \times M^n - A \bmod \Delta)$ is zero.

Next we will show that the "relative" 2-fold covering map

$$(M^n \times M^n - A, \Delta) \xrightarrow{c} (M^n * M^n - A', \Delta')$$

has a Gysin cohomology sequence. The difficulty lies in the fact that this

map is a covering only in the complement of the diagonal Δ. If U denotes any symmetric neighborhood of Δ, and U' denotes its image in the collapsed space, consider the diagram

$$\begin{tikzcd}
(M^n \times M^n - A - \Delta, U - \Delta) \arrow[r, "e"] \arrow[d, "c_1"] & (M^n \times M^n - A, U) \arrow[d, "c_2"] & (M^n \times M^n - A, \Delta) \arrow[l] \arrow[d, "c"] \\
(M^n * M^n - A' - \Delta', U' - \Delta') \arrow[r, "e'"] & (M^n * M^n - A', U') & (M^n * M^n - A', \Delta'). \arrow[l]
\end{tikzcd}$$

Since c_1 is a true 2-fold covering, it has a Gysin sequence. Since e and e' are excisions, the map c_2 also has a Gysin sequence. But since Δ and $M^n \times M^n - A$ are both absolute neighborhood retracts, the cohomology group $H^*(M^n \times M^n - A \bmod \Delta)$ is the direct limit of the groups $H^*(M^n \times M^n - A \bmod U)$ as U ranges over all symmetric neighborhoods. A similar description holds for $H^*(M^n * M^n - A' \bmod \Delta')$. Since the direct limit of exact sequences is again an exact sequence, this gives us the required Gysin sequence

$$\to H^k(M^n \times M^n - A, \Delta) \to H^k(M^n * M^n - A', \Delta')$$
$$\to H^{k+1}(M^n * M^n - A', \Delta') \xrightarrow{c^*} H^{k+1}(M^n \times M^n - A, \Delta) \to.$$

Since the left and right hand groups in this sequence are known to be zero, induction shows that the middle groups are also zero. Therefore the corresponding homology groups $H_k(M^n * M^n - A' \bmod \Delta')$ are zero. This implies that the inclusion homomorphism $H_*(\Delta') \to H_*(M^n * M^n - A')$ is an isomorphism. It has already been shown that the inclusion homomorphism $H_*(\Delta) \to H_*(M^n \times M^n - A)$ is an isomorphism. Since the map $i_2 | \Delta : \Delta \to \Delta'$ is a homeomorphism, this implies that

$$i_{2*} : H_*(M^n \times M^n - A) \to H_*(M^n * M^n - A')$$

is an isomorphism; which completes the proof of Lemma 1.

Consider the diagram

$$\begin{tikzcd}
 & H_n(M^n) \\
H_n(M^n) \arrow[ur, "f_*"] \arrow[r, "f_{1*}"] \arrow[d, "i_{1*}"] & H_n(M^n \times M^n - A) \arrow[u, "p_{1*}"'] \arrow[d, "i_{2*}"] \\
H_n(M^n/T) \arrow[r, "f_{2*}"] & H_n(M^n * M^n - A').
\end{tikzcd}$$

It is clear that $i_{1*} = 0$. Since i_{2*} is an isomorphism, this implies that $f_{1*} = 0$, and hence $f_* = 0$. This contradicts the hypothesis of Theorem 1, and completes the proof.

2. A survey of the problem. We will say that a group G has *property I* if every abelian subgroup is cyclic. This is equivalent to the condition that every p-subgroup is either cyclic or, for $p = 2$, a generalized quaternion group.[2]

According to P. Smith [5], every group which acts on a homology sphere without fixed points has property I. This result can be given more precision by noting that if G acts on S^n without fixed points then[2] the cohomology ring $H^*(G, Z)$ has period $n+1$. According to Artin and Tate a group has property I if and only if it has periodic cohomology.[2]

An enumeration of all groups having property I has been given by Zassenhaus ([8] Theorem 7) in the solvable case, and by M. Suzuki ([6] Theorem E) in the non-solvable case.

Corollary 1 gives a further restriction which the group G must satisfy. (For example the non-abelian group of order $2p$ has property I and yet cannot act on a sphere without fixed points.) However it seems likely that more restrictions are still needed.

If G is required to act orthogonally on S^n, then the problem has been solved by Zassenhaus. ([8] Theorems 3, 8, 11, 16. See also Vincent [7].) At present, every group which is known to act on S^n without fixed points can actually act orthogonally.

The following two results of Zassenhaus are especially striking. We will say that G has *property II* if every subgroup of order pq is cyclic. Here p and q range over not necessarily distinct primes.

(a) Every group which acts orthogonally on S^n without fixed points has property II.

(b) Every solvable group with property II can act orthogonally on some S^n without fixed points.

Evidently the major unsolved problem is the following. Does every group which acts on S^n without fixed points have property II? Smith's result implies that every subgroup of order p^2 is cyclic, and Corollary 1 implies that every subgroup of order $2p$ is cyclic. The first unsolved case is therefore the following. Can the non-abelian group of order 21 act on a sphere without fixed points?

[2] See Cartan and Eilenberg [1], Chapters XII, Section 11 and XVI, Section 9.

3. The case $n = 3$. The main object of this section will be to summarize the gap in our knowledge concerning groups which act on the 3-sphere without fixed points (or concerning finite fundamental groups of closed 3-manifolds). This will be done by enumerating those groups which satisfy the known restrictions, but which cannot act orthogonally on S^3 without fixed points. (Theorem 3.) It is interesting to note that the groups in question are all solvable, and are all known to act on higher dimensional spheres without fixed points. The proofs are routine, and will be largely omitted.

The sphere S^3 is itself a topological group, isomorphic to the two-fold covering space of the rotation group $SO(3)$. Let $h: S^3 \to SO(3)$ denote the covering. Then each finite subgroup G of $SO(3)$ determines a subgroup $h^{-1}(G)$ of S^3 with twice as many elements. Corresponding to the dihedral group with $2n$ elements, $n \geqq 2$, we obtain a subgroup Q_{4n} of S^3 with presentation $(x, y : x^2 = (xy)^2 = y^n)$. (If n is a power of 2, then Q_{4n} is the generalized quaternion group.) Corresponding to the tetrahedral, octahedral, and icosohedral groups, we obtain subgroups P_{24}, P_{48} and P_{120} of S^3 with presentations $(x, y : x^2 = (xy)^3 = y^n, x^4 = 1)$, where $n = 3, 4, 5$ respectively.

The next theorem is essentially due to Hopf [2]. For a detailed study of the quotient manifolds see Seifert and Threlfall [4].

THEOREM 2. *The following is a list of all finite groups which act orthogonally on S^3 without fixed points.*

1) *The groups* 1, Q_{8n}, P_{48}, *and* P_{120}.

2) *The groups* $D_{2^k(2n+1)}$ *with presentation*

$$(x, y : x^{2^k} = 1,\ y^{2n+1} = 1,\ xyx^{-1} = y^{-1}) \text{ where } k \geqq 2, n \geqq 1.$$

3) *The groups* $P'_{8\cdot 3^k}$ *with presentation*

$$(x, y, z : x^2 = (xy)^2 = y^2, zxz^{-1} = y, zyz^{-1} = xy, z^{3^k} = 1) \text{ where } k \geqq 1.$$

4) *The direct product of any of these groups with a cyclic group of relatively prime order.*

(Note that the groups Q_{4n} with n odd are isomorphic to D_{4n}. The group P_{24} is isomorphic to P'_{24}.)

For a sketch of the proof we refer to Hopf's paper.

Given relatively prime positive integers $8n$, k, l let $Q(8n, k, l)$ denote the group with presentation

$$(x, y, z : x^2 = (xy)^2 = y^{2n}, z^{kl} = 1, xzx^{-1} = z^r, yzy^{-1} = z^{-1}),$$

where $r \equiv -1 \pmod{k}$, $r \equiv +1 \pmod{l}$. Thus for $k = l = 1$ we have the group Q_{8n}. Note that $Q(8n, k, l)$ is isomorphic to $Q(8n, 1, k)$. For n odd, the three integers n, k, l may be permuted without changing the group.

For any odd integer r let P''_{48r} denote the group extension with subgroup Z_r and factor group P_{48} such that the 3-Sylow subgroup is cyclic, and such that P_{48} acts on Z_r as follows. Elements in the commutator subgroup P_{24} act trivially, while the remaining elements of P_{48} carry each element of Z_r into its inverse.

THEOREM 3. *The following is a list of all finite groups which* (a) *have cohomology of period* 4, (b) *have at most one element of order* 2, *and* (c) *are not listed in Theorem* 2.

1) *The groups* $Q(8n, k, l)$ *with* $8n, k, l$ *pairwise relatively prime and either* a) n *odd and* $n > k > l \geqq 1$ *or* b) n *even and* $n \geqq 2$, $k > l \geqq 1$.

2) *The groups* P''_{48r} *with* r *odd*, $r \geqq 3$.

3) *The product of any of these groups with a cyclic group of relatively prime order.*

Thus the smallest group which may or may not be the fundamental group of a 3-manifold is $Q(16, 3, 1)$ with 48 elements.

A few of these groups can be eliminated by a different argument:

LEMMA 2. *If r is not a power of* 3, *then* P''_{48r} *cannot be the fundamental group of a 3-manifold.*

Proof. Let $r = 3^{k-1}s$ with $(s, 6) = 1$, $s \geqq 5$. Then P''_{48r} contains a normal subgroup isomorphic to $P'_{8 \cdot 3^k}$, the factor group being isomorphic to D_{2s}.

Now given a manifold M^3 with $\pi_1(M^3) \approx P''_{48r}$ let $M_1{}^3$ denote the $2s$-fold covering with fundamental group $\pi_1(M_1{}^3) = P'_{8 \cdot 3^k}$. The first homology group of $M_1{}^3$ is the abelianized group, which turns out to be Z_3. Therefore $M_1{}^3$ has the mod two homology of the 3-sphere. But the group D_{2s} of covering transformations acts on $M_1{}^3$ without fixed points. Since D_{2s} has an element of order two which is not in the center, this contradicts Corollary 1, and completes the proof.

The proof of Theorem 3 consists in going over the groups listed by Zassenhaus and Suzuki, eliminating those which do not have cohomology of period four. The spectral sequence of a group extension is the main tool needed.

PRINCETON UNIVERSITY.

REFERENCES.

[1] H. Cartan and S. Eilenberg, *Homological Algebra*, Princeton, 1956.

[2] H. Hopf, "Zum Clifford-Kleinschen Raumproblem," *Mathematische Annalen*, vol. 95 (1925-26), pp. 313-319.

[3] P. Olum, "Mappings of manifolds and the notion of degree," *Annals of Mathematics*, vol. 58 (1953), pp. 458-480.

[4] H. Seifert and W. Threlfall, "Topologische Untersuchung der Diskontinuitätsbereiche endlicher Bewegungsgruppen des dreidimensionalen sphärischen Raumes," *Mathematische Annalen*, vol. 104 (1930-31), pp. 1-70.

[5] P. Smith, "Permutable periodic transformations," *Proceedings of the National Academy of Sciences*, vol. 30 (1944), pp. 105-108.

[6] M. Suzuki, "On finite groups with cyclic Sylow subgroups for all odd primes," *American Journal of Mathematics*, vol. 77 (1955), pp. 657-691.

[7] G. Vincent, "Les groupes linéaires finis san points fixes," *Commentarii Mathematici Helvetici*, vol. 20 (1947), pp. 117-171.

[8] H. Zassenhaus, "Über endliche Fastkörper," *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, vol. 11 (1936), pp. 187-220.

ON A CLASS OF TRANSFORMATION GROUPS.*

By Andrew M. Gleason and Richard S. Palais.

In order to apply our rather deep understanding of the structure of Lie groups to the study of transformation groups it is natural to try to single out a class of transformation groups which are in some sense naturally Lie groups. In this paper we introduce such a class and commence their study.

In Section 1 the notion of a Lie transformation group is introduced. Roughly, these are groups H of homeomorphisms of a space X which admit a Lie group topology which is strong enough to make the evaluation mapping $(h,x) \to h(x)$ of $H \times X$ into X continuous, yet weak enough so that H gets all the one-parameter subgroups it deserves by virtue of the way it acts on X (see the definition of admissibly weak below). Such a topology is uniquely determined if it exists and our efforts are in the main concerned with the question of when it exists and how one may effectively put one's hands on it when it does. A natural candidate for this so-called Lie topology is of course the compact-open topology for H. However, if one considers the example of a dense one-parameter subgroup H of the torus X acting on X by translation, it appears that this is not the general answer. In this example if we modify the compact-open topology by adding to the open sets all their arc components (getting in this way what we call the modified compact-open topology), we get the Lie topology of H. That this is a fairly general fact is one of our main results (Theorem 5.14). The latter theorem moreover shows that the reason that the compact-open topology was not good enough in the above example is connected with the fact that H was not closed in the group of all homeomorphisms of X, relative to the compact-open topology. Theorem 5.14 also states that for a large class of interesting cases the weakness condition for a Lie topology is redundant.

The remainder of the paper is concerned with developing a certain criterion for deciding when a topological group is a Lie group and applying this criterion to derive a general necessary and sufficient condition for groups of homeomorphisms of locally compact, locally connected finite dimensional metric spaces to be Lie transformation groups. The criterion is remarkable in that local compactness is not one of the assumptions. It states in fact

* Received February 11, 1957.

that a locally arcwise connected topological group is a Lie group provided that its compact metrizable subspaces are of bounded dimension.

1. Lie transformation groups. Let G be a topological group and X a topological space. By an *action* of G on X we mean a homomorphism $\phi\colon g \to \phi_g$ of G into the group of homeomorphisms of X such that the map $(g,x) \to \phi_g(x)$ of $G \times X$ into X is continuous. If H is a group of homeomorphisms of X, then a topology for H will be called *admissibly strong* if it renders the map $(h,x) \to h(x)$ of $H \times X \to X$ continuous. We note that we do *not* demand of an admissibly strong topology that it make H a topological group; however if H is a topological group in a given topology, then clearly that topology is admissibly strong if and only if it makes the identity map of H on itself an action of H on X. Moreover if we denote by R the additive group of real numbers then:

1.1. PROPOSITION. *Let H be a topological group whose underlying group is a group of homeomorphisms of a space X. If the topology of H is admissibly strong, then each one-parameter subgroup of H is an action of R on X.*

We shall call a topology for a group H of homeomorphisms of a space X *admissibly weak* if every action of R on X whose range is in H is a continuous map of R into H with respect to this topology. Again we note that an admissibly weak topology for H is not required to make H a topological group. However from 1.1 and the definition of admissibly weak we clearly have:

1.2. PROPOSITION. *Let H be a topological group whose underlying group is a group of homeomorphisms of a space X. If the topology for H is both admissibly weak and admissibly strong then the one-parameter subgroups of H are exactly the actions of R on X whose ranges lie in H.*

The terminology 'admissibly strong' and 'admissibly weak' is justified by the following trivial observation.

1.3. PROPOSITION. *Let H be a group of homeomorphisms of a space X. A topology for H which is stronger (weaker) than an admissibly strong (weak) topology is itself admissibly strong (weak).*

Some authors use the term admissible for topologies that we call admissibly strong. For this reason we shall not succumb to the temptation of calling admissible those topologies which are at once admissibly strong and admissibly weak.

1.4. *Definition.* Let H be a group of homeomorphisms of a space X. A *Lie topology* for H is a topology for H which is both admissiblly strong and admissibly weak and which furthermore makes H a Lie group.

The following well-known fact is an immediate consequence of the existence of canonical coordinate systems of the second kind in Lie groups.

1.5. LEMMA. *Let G and H be Lie groups and h a homomorphism of the underlying group of G into the underlying group of H. A necessary and sufficient condition for h to be continuous is that $h \circ \phi$ be a one parameter subgroup of H whenever ϕ is a one-parameter subgroup of G. In particular, if G and H have the same underlying group and the same one-parameter subgroups, they are identical.*

The following proposition follows directly from 1,2, 1.5, and the definition of a Lie topology.

1.6. PROPOSITION. *A group of homeomorphisms of a topological space admits at most one Lie topology.*

1.7. *Definition.* A group of homeomorphisms of a topological space X will be called a *Lie transformation group of X* if it admits a Lie topology.

The unique Lie topology for a Lie transformation group G will be called the Lie topology for G and properties meaningful for a Lie group when used in reference to G are to be interpreted relative to its Lie topology.

2. A theorem on arcwise connected spaces. A theorem somewhat more general than the next lemma is proved on page 115 of [6], and a still more general result is indicated in exercise 7, page 80 of [8].

2.1. LEMMA. *A partitioning of the unit interval into at most countably many disjoint closed sets is trivial, i.e., contains only one element.*

2.2. THEOREM. *A partitioning of an arcwise connected space X into at most countably many disjoint closed sets is trivial.*

Proof. Let $\{E_n\}$ be such a partitioning and let $p, q \in X$. We must show that p and q are in the same E_n. Let f be a continuous map of the unit interval into X such that $f(0) = p$ and $f(1) = q$. Applying the lemma to the partitioning $\{f^{-1}(E_n)\}$ of the unit interval we see that for some n $f^{-1}(E_n)$ is the entire unit interval. Hence $p = f(0)$ and $q = f(1)$ belong to E_n.

3. Making a topology locally arcwise connected. Most, if not all, of the results of this section are known, but they belong to the realm of folk-theorems and are apparently not easily available in the literature.

Let $(X, \mathcal{T})$ be a topological space (i.e., X a set and $\mathcal{T}$ a topology for X) and let $\mathfrak{B}$ be the set of arc components of all open subspaces of $(X, \mathcal{T})$. Suppose B_1 and B_2 are elements of $\mathfrak{B}$ and let B_i be an arc component of $\mathfrak{O}_i \in \mathcal{T}$. Then if $p \in B_1 \cap B_2$, the arc component of p in $\mathfrak{O}_1 \cap \mathfrak{O}_2$, which belongs to $\mathfrak{B}$, is clearly a subset of $B_1 \cap B_2$. Thus $B_1 \cap B_2$ is a union of sets from $\mathfrak{B}$ and hence $\mathfrak{B}$ is a base for a new topology $\mathcal{M}(\mathcal{T})$ for X which is clearly stronger than $\mathcal{T}$.

3.1. *Definition.* We define an operation $\mathcal{M}$ on topologies as follows: if $\mathcal{T}$ is a topology for a set X then $\mathcal{M}(\mathcal{T})$ is the topology for X which has as a base all arc components of open subspaces of $(X, \mathcal{T})$.

The following theorem summarizes some of the most important properties of the operation $\mathcal{M}$.

3.2. THEOREM. *Let $(X, \mathcal{T})$ be a topological space.*

(1) *If $\mathcal{T}$ satisfies the first axiom of countability, so does $\mathcal{M}(\mathcal{T})$.*

(2) *If Z is a locally arcwise connected space and f is a function from Z into X continuous relative to the topology $\mathcal{T}$, then f is also continuous relative to $\mathcal{M}(\mathcal{T})$. In particular $(X, \mathcal{T})$ and $(X, \mathcal{M}(\mathcal{T}))$ have the same arcs.*

(3) *$(X, \mathcal{M}(\mathcal{T}))$ is locally arcwise connected, and in fact $\mathcal{M}(\mathcal{T})$ can be characterized as the weakest locally arcwise connected topology for X which is stronger than $\mathcal{T}$. Hence $\mathcal{M}$ is idempotent.*

(4) *The components of an open subset of $(X, \mathcal{M}(\mathcal{T}))$ are just its arc components when regarded as a subspace of $(X, \mathcal{T})$. In particular the components of $(X, \mathcal{M}(\mathcal{T}))$ are the arc components of $(X, \mathcal{T})$.*

(5) *If X is a group and $(X, \mathcal{T})$ a topological group, then $(X, \mathcal{M}(\mathcal{T}))$ is also a topological group and it has the same one-parameter subgroups as $(X, \mathcal{T})$.*

Proof. Given $x \in X$ and a countable base $\{\mathfrak{O}_n\}$ for the $\mathcal{T}$-neighborhoods of x we get a countable base $\{\mathfrak{O}_n'\}$ for the $\mathcal{M}(\mathcal{T})$-neighborhoods of x by taking $\mathfrak{O}_n'$ to be the arc component of x in $\mathfrak{O}_n$ (relative to the topology $\mathcal{T}$, of course). This proves (1).

Let $\mathfrak{B}$ be the set of all arc components of open subspaces of $(X, \mathcal{T})$ so that by definition $\mathfrak{B}$ is a base for $\mathfrak{M}(\mathcal{T})$.

Suppose f is a function from the locally arcwise connected space Z into X which is continuous relative to $\mathcal{T}$. Given $B \in \mathfrak{B}$ we will show that $f^{-1}(B)$ is open in Z which will prove (2). By definition of $\mathfrak{B}$ we can choose $\mathfrak{O} \in \mathcal{T}$ such that B is an arc component of $\mathfrak{O}$. Given $p \in f^{-1}(B)$ let W be the arc component of p in $f^{-1}(\mathfrak{O})$. Then $f(W)$ is arcwise connected, included in $\mathfrak{O}$, and meets B at $f(p)$; hence $f(W) \subseteq B$. Since Z is locally arcwise connected and $f^{-1}(\mathfrak{O})$ is open, W is open. Thus a neighborhood of p is included in $f^{-1}(B)$ so $f^{-1}(B)$ is open.

Next let $B \in \mathfrak{B}$. By definition of $\mathfrak{B}$, B is arcwise connected when regarded as a subspace of $(X, \mathcal{T})$. Hence by (2) B is an arcwise connected subspace of $(X, \mathfrak{M}(\mathcal{T}))$. Thus $\mathfrak{M}(\mathcal{T})$ has a base consisting of arcwise connected sets so, by definition, $\mathfrak{M}(\mathcal{T})$ is locally arcwise connected. Since every $\mathfrak{O} \in \mathcal{T}$ is the union of its arc components and hence belongs to $\mathfrak{M}(\mathcal{T})$ it follows that $\mathfrak{M}(\mathcal{T})$ is stronger than $\mathcal{T}$. Suppose $\mathcal{T}'$ is a locally arcwise topology for X stronger than $\mathcal{T}$. Then the identity mapping f of $Z = (X, \mathcal{T}')$ into $(X, \mathcal{T})$ is continuous and hence, by (2), f is a continuous map of Z into $(X, \mathfrak{M}(\mathcal{T}))$, i.e., $\mathcal{T}'$ is stronger that $\mathfrak{M}(\mathcal{T})$. This proves (3).

If $\mathfrak{O}$ is an open subspace of $(X, \mathfrak{M}(\mathcal{T}))$ then, since $\mathfrak{M}(\mathcal{T})$ is locally arcwise connected, the components of $\mathfrak{O}$ are the same as the arc components of $\mathfrak{O}$. On the other hand, by (2), the arc components of $\mathfrak{O}$ are the same whether $\mathfrak{O}$ is regarded as a subspace of $(X, \mathcal{T})$ or $(X, \mathfrak{M}(\mathcal{T}))$. This proves (4).

Finally, suppose that X is a group and let f be the map $(x, y) \to xy^{-1}$ of $X \times X \to X$. If $(X, \mathcal{T})$ is a topological group then f is a continuous map of $(X, \mathcal{T}) \times (X, \mathcal{T}) \to (X, \mathcal{T})$ and *a fortiori* (since $\mathfrak{M}(\mathcal{T})$ is stronger than $\mathcal{T}$) f is a continuous map of $(X, \mathfrak{M}(\mathcal{T})) \times (X, \mathfrak{M}(\mathcal{T})) \to (X, \mathcal{T})$. Since by (3) $(X, \mathfrak{M}(\mathcal{T})) \times (X, \mathfrak{M}(\mathcal{T}))$ is locally arcwise connected, it follows from (2) that f is a continuous map of

$$(X, \mathfrak{M}(\mathcal{T})) \times (X, \mathfrak{M}(\mathcal{T})) \to (X, \mathfrak{M}(\mathcal{T})),$$

i.e., that $(X, \mathfrak{M}(\mathcal{T}))$ is a topological group. It also follows from (2) that $(X, \mathcal{T})$ and $(X, \mathfrak{M}(\mathcal{T}))$ have the same arcs and hence the same one-parameter subgroups. This proves (5).

3.3. *Definition.* If $\mathfrak{G} = (G, \mathcal{T})$ is a topological group, then we call $(G, \mathfrak{M}(\mathcal{T}))$ the *associated locally arcwise connected group of* $\mathfrak{G}$.

4. Weakening the topology of a Lie group.

4.1. THEOREM. *Let ϕ be a one-to-one representation of a locally compact group G satisfying the second axiom of countability onto a locally arcwise connected group H. Then ϕ^{-1} is continuous, i.e., ϕ is an isomorphism of G with H.*

Proof. Let V be a compact neighborhood of e_G, the identity of G. It will suffice to show that $\phi(V)$ is a neighborhood of e_H, the identity of H. Choose an open, symmetric neighborhood U of e_G such that $\bar{U}^2 \subseteq V$. Then $V-U$ is compact, so $\phi(V-U)$ is compact and hence that complement of $\phi(V-U)$ is a neighborhood of e_H. Let X be an arcwise connected neighborhood of e_H such that XX^{-1} does not meet $\phi(V-U)$.

Given g_1 and g_2 in $\phi^{-1}(X)$ we put $g_1 \sim g_2$ if and only in $g_1g_2^{-1} \in \bar{U}$. Since $\bar{U}$ is a symmetric neighborhood of e_G, it follows that $\sim$ is a symmetric, reflexive relation on $\phi^{-1}(X)$. If $g_1 \sim g_2$ and $g_2 \sim g_3$ then $g_1g_3^{-1} = (g_1g_2^{-1})(g_2g_3^{-1}) \in \bar{U}^2 \subseteq V$. But $\phi(g_1g_3^{-1}) = \phi(g_1)\phi(g_3)^{-1} \in XX^{-1}$ and since XX^{-1} is disjoint from $\phi(V-U)$, it follows that $g_1g_3^{-1} \in U \subseteq \bar{U}$ so $g_1 \sim g_3$. Hence $\sim$ is also transitive and hence is an equivalence relation on $\phi^{-1}(X)$. Let $\{g_\alpha\}$ be a complete set of representatives of $\phi^{-1}(X)$ under $\sim$, one of which we can take to be e_G. Given $g \in \phi^{-1}(X)$ we can find a g_α such that $g_\alpha \sim g$ so $g \in \bar{U}g_\alpha$. Thus $\{\bar{U}g_\alpha\}$ is a covering of $\phi^{-1}(X)$. If $g \in \bar{U}g_\alpha \cap \bar{U}g_\beta$, then $g_\alpha g^{-1} \in \bar{U}^{-1} = \bar{U}$ and $gg_\beta^{-1} \in \bar{U}$, so $g_\alpha g_\beta^{-1} \in \bar{U}^2 \subseteq V$. But $\phi(g_\alpha g_\beta^{-1}) \in XX^{-1}$ which is disjoint from $\phi(V-U)$ so $g_\alpha g_\beta^{-1} \in U \subseteq \bar{U}$ so $g_\alpha \sim g_\beta$ and $\alpha = \beta$. Hence the $\bar{U}g_\alpha$ are disjoint and therefore, since they have non-empty interiors and G satisfies the second axiom of countability, it follows that $\{\bar{U}g_\alpha\}$ is a countable set. Now since ϕ is one-to-one, $\{X \cap \phi(\bar{U}g_\alpha)\}$ is a countable disjoint covering of X. Moreover since $\bar{U}$ is closed and included in V it is compact. Hence each $\bar{U}g_\alpha$ is compact, so each $\phi(\bar{U}g_\alpha)$ is compact, so each $X \cap \phi(\bar{U}g_\alpha)$ is closed in X. Now $X \cap \phi(\bar{U}e_G)$ is not empty, and in fact contains e_H. Since X is arcwise connected it follows from (2.2) that $X \cap \phi(\bar{U}) = X$. Thus $X \subseteq \phi(\bar{U}) \subseteq \phi(\bar{U}^2) \subseteq \phi(V)$ so $\phi(V)$ is a neighborhood of e_H as was to be proved.

4.2. THEOREM. *Let $\mathfrak{G}$ be a locally arcwise connected, locally compact group satisfying the second axiom of countability. If the underlying group of $\mathfrak{G}$ is a topological group $\mathfrak{G}^*$ in a topology weaker than the topology of $\mathfrak{G}$, then $\mathfrak{G}$ is the associated locally arcwise connected group of $\mathfrak{G}^*$. In particular, $\mathfrak{G}^*$ is locally arcwise connected, then $\mathfrak{G} = \mathfrak{G}^*$. In any case the arc components of open subspaces of $\mathfrak{G}^*$ form a base for the topology of $\mathfrak{G}$ and both $\mathfrak{G}$ and $\mathfrak{G}^*$ have the same one-parameter subgroups.*

Proof. Let $\mathfrak{G}^{**}$ be the associated locally arcwise connected group of $\mathfrak{G}^*$. Since, by definition, the topology of $\mathfrak{G}^{**}$ is the weakest locally arcwise connected topology stronger than the topology of $\mathfrak{G}^*$, it follows that the topology of $\mathfrak{G}$ is stronger than the topology of $\mathfrak{G}^{**}$, i.e., the identity map ϕ is a representation of $\mathfrak{G}$ on $\mathfrak{G}^{**}$. It follows from (4.1) that ϕ is an isomorphism of $\mathfrak{G}$ with $\mathfrak{G}^{**}$, i.e. $\mathfrak{G} = \mathfrak{G}^{**}$.

4.3. COROLLARY. *If $\mathfrak{G} = (G, \mathfrak{T})$ is a Lie group satisfying the second axiom of countability, then the topology $\mathfrak{T}$ of $\mathfrak{G}$ is minimal in the set of all locally arcwise connected group topologies for G.*

5. The compact-open and modified compact-open topologies. Let X be a topological space and let $\mathfrak{H}(X)$ denote the group of all homeomorphisms of X on itself. Given subsets of X $K_1, \cdots, K_n$ and $\mathfrak{O}_1, \cdots, \mathfrak{O}_n$ with the K_i compact and the $\mathfrak{O}_i$ open define

$$(K_1, \cdots, K_n; \mathfrak{O}_1, \cdots, \mathfrak{O}_n) = \{h \in \mathfrak{H}(X) : h(K_i) \subseteq \mathfrak{O}_i, i = 1, \cdots, n\}.$$

The compact-open topology for $\mathfrak{H}(X)$ is by definition the topology in which sets of the above form are a basis. If H is a subgroup of $\mathfrak{H}(X)$, then the compact-open topology for H is the topology induced on H by the compact-open topology for $\mathfrak{H}(X)$; equivalently it is the topology which has as a basis sets of the form

$$(K_1, \cdots, K_n; \mathfrak{O}_1, \cdots, \mathfrak{O}_n)_H = \{h \in H : h(K_i) \subseteq \mathfrak{O}_i, i = 1, \cdots, n\}.$$

We refer the reader to [1] for details concerning the compact-open topology (it is called the k-topology there). We will need the following facts proved in [1].

5.1. *If X is locally compact, then the compact-open topology for a group of homeomorphisms of X is admissibly strong and is weaker than any other admissibly strong topology.*

5.2. *If X is locally compact and locally connected, then every group of homeomorphisms of X is a topological group in its compact-open topology.*

Immediate from the definition of the compact-open topology is

5.3. PROPOSITION. *If H is a group of homeomorphisms of a space and G a subgroup of H, then the compact-open topology for H induces on G the compact-open topology for G.*

Another fact we will need is

5.4. PROPOSITION. *If X is a locally compact space satisfying the second axiom of countability, then the compact-open topology for any group H of homeomorphisms of X also satisfies the second axiom of countability.*

Proof. Choose a basis for the topology of X consisting of a sequence $\{\mathcal{O}_i\}$ such that each $\bar{\mathcal{O}}_i$ is compact. Then sets of the form $(\bar{\mathcal{O}}_{i_1}, \cdots, \bar{\mathcal{O}}_{i_n}; \mathcal{O}_{j_1}, \cdots, \mathcal{O}_{j_n})_H$ give a countable base for the compact-open topology for H.

If $\mathcal{G}$ is a topological group, then the bilateral uniform structure for $\mathcal{G}$ is that uniform structure generated by uniformities of the form $\{(g,h) \in \mathcal{G} \times \mathcal{G} : gh^{-1}$ and $g^{-1}h \in V\}$ for some neighborhood V of the indentity in $\mathcal{G}$. Like the left and right uniform structures for $\mathcal{G}$ the bilateral uniform structure is compatible with the topology of $\mathcal{G}$. It has a countable base, and is hence equivalent to a metric, if and only if $\mathcal{G}$ satisfies the first axiom of countability. Now in [1] Arens shows that if X is a locally compact, locally connected space and $\mathcal{H}(X)$ is the group of all homeomorphisms of X made into a topological group (5.2) by giving it its compact-open topology, then $\mathcal{H}(X)$ is complete in its bilateral uniform structure (but not generally in its left and right uniform structures). If we now assume that X satisfies the second axiom of countability and use (5.4), we get a fact mentioned in a footnote of [1].

5.5. PROPOSITION. *Let X be a locally compact, locally connected space satisfying the second axiom of countability and let H be a group of homeomorphisms of X which is closed, relative to the compact-open topology, in the group of all homeomorphism of X. Then the compact-open topology for H can be derived from a complete metric, hence H is of the second category in its compact-open topology.*

Now it is a well-known fact that a continuous one-to-one homomorphism of a locally compact topological group G satisfying the second axiom of countability onto a topological group H of the second category is necessarily bicontinuous (see, for example, Theorem XIII, page 65 of Pontrjagin's *Topological Groups,* where the proof is given under the assumption that H is locally compact, but only the consequence, that H is of the second category, is actually used). Using this result together with (5.1) and (5.5) we get:

5.6. PROPOSITION. *Let X be a locally compact, locally connected space satisfying the second axiom of countability and let H be a group of homeomorphisms of X which is closed, relative to the compact-open topology, in the group of all homeomorphisms of X. If H is a topological group in an admissibly strong, locally compact topology $\mathfrak{T}$ which satisfies the second axiom of countability, then $\mathfrak{T}$ is the compact-open topology for H.*

5.7. PROPOSITION. *Let X be a locally compact space, G a topological group, and ϕ an action of G on X whose range lies in a group H of homeomorphisms of X. Then ϕ is a continuous map of G into H when the latter is given its compact-open topology.*

Proof. Let K be the kernel of ϕ. It follows from the fact that ϕ is an action that K is closed in G. Let h be the canonical homomorphism of G on G/K and $\phi = \bar{\phi} \circ h$ the canonical factoring of ϕ. Since h is continuous, it will suffice to show that $\bar{\phi}$ is continuous when H is given its compact-open topology. Now it follows from the fact that h is an open mapping that $\bar{\phi}$ is an action of G/K on X and of course $\bar{\phi}$ is one-to-one. Thus it suffices to prove the theorem when ϕ is one-to-one (i.e., effective in the usual terminology). It is then no loss of generality to assume that G is a subgroup of H and that ϕ is the injection mapping. We can then restate the proposition as follows: if the topology of G is admissibly strong it is stronger than the topology induced on G by the compact-open topology of H. Since by (5.3) H induces the compact-open topology on G, this restatement is a consequence of (5.1).

Taking $G = R$ in (5.7) and recalling the definition of admissibly weak we have

5.8. COROLLARY. *If X is a locally compact space, then the compact-open topology for a group of homeomorphisms of X is always admissibly weak.*

5.9. *Definition.* Let G be a group of homeomorphisms of a space X. the *modified compact-open topology* for G is the topology resulting from applying the operation $\mathfrak{M}$ (3.1) to the compact-open topology for G. In other words it is the topology for G in which the arc components of open subspaces of G (relative to the compact-open topology) form a base.

A word of caution: since the operation $\mathfrak{M}$ does not in general commute with the operation of inducing a topology on a subspace, there is no analogue of (5.3) for the modified compact-open topology, i.e., the modified compact-open topology for a group H of homeomorphisms des not in general induce on a subgroup G of H the modified compact-open topology for G.

5.10. PROPOSITION. *If X is a locally compact space and G is a group of homeomorphisms of X, then the modified compact-open topology for G can be characterized as the weakest admissibly strong topology for G which is locally arcwise connected. Moreover, the modified compact-open topology for G is also admissibly weak.*

Proof. The first conclusion follows from (5.1) and (3.2(3)). If ϕ is an action of R on X with range in G, then by (5.7) ϕ is continuous when G is given its compact-open topology and hence by (3.2(2)) when G is given its modified compact-open topology, i.e., the modified compact-open topology for G is admissibly weak.

5.11. PROPOSITION. *Let X be a locally compact space and G a group of homeomorphisms of X. Then the modified compact-open topology is a Lie topology for G if and only if it makes G a Lie group. The same is true of the compact-open topology.*

Proof. This follows directly from the definition of a Lie topology (1.4) since we have seen (5.1, 5.8, 5.10) that both the compact-open and modified compact-open topologies are admissibly strong and admissibly weak.

5.12. PROPOSITION. *Let G be a group of homeomorphisms of a locally compact, locally connected space X. Then G is a locally arcwise connected topological group in its modified compact-open topology. Moreover, G has the same one-parameter subgroups when given either its compact-open topology or its modified compact-open topology, in fact, in each case they are exactly the actions of R on X with range in G.*

Proof. The first conclusion follows from (5.2) and (3.2(5)). Since we have seen that both the compact-open and modified compact-open topologies are admissibly weak and admissibly strong, the final conclusion follows from (1.2).

5.13. PROPOSITION. *If X is a locally compact space satisfying the second axiom of countability, then the modified compact-open topology for any group of homeomorphisms of X satisfies the first axiom of countability.*

Proof. (5.4) and (3.2(1)).

In general, however, it will not be true in the above case that the modified compact-open topology, like the compact-open topology, satisfies the second axiom of countability.

The following is one of our main results concerning Lie transformation groups.

5.14. THEOREM. *Let $\mathfrak{G} = (G, \mathfrak{T})$ be a Lie group satisfying the second axiom of countability whose underlying group G is a group of homeomorphisms of a locally compact, locally connected space X. If the topology $\mathfrak{T}$ of $\mathfrak{G}$ is admissibly strong, then it is automatically admissibly weak and hence G is a Lie transformation group of X and $\mathfrak{T}$ its Lie topology. Moreover $\mathfrak{T}$*

is the modified compact-open topology for G and if X satisfies the second axiom of countability and G is closed, relative to the compact-open topology, in the group of all homeomorphisms of X, then $\mathfrak{T}$ is the compact-open topology for G.

Proof. Since by (5.10) the modified compact-open topology for G is admissibly weak, everything will follow once we show that $\mathfrak{T}$ is the modified compact-open topology (the final conclusion is a consequence of (5.6)).

Let $\mathfrak{G}^*$ denote G taken with the modified compact-open topology. By (5.12) $\mathfrak{G}^*$ is a locally arcwise connected topological group. Since $\mathfrak{G}$ is locally arcwise connected and has an admissibly strong topology, it follows from (5.10) that the topology of $\mathfrak{G}$ is stronger than the topology of $\mathfrak{G}^*$. Then by (4.3) $\mathfrak{G} = \mathfrak{G}^*$, i.e., $\mathfrak{T}$ is the modified compact-open topology for G.

5.15. Corollary. *If a Lie topology for a group G of homeomorphisms of a locally compact, locally connected space satisfies the second axiom of countability, then it is the modified compact-open topology for G.*

5.16. Corollary. *If X is a locally compact, locally connected space, then the connected Lie transformation groups of X are precisely the groups of homeomorphisms of X which are connected Lie groups in their modified compact-open topology.*

Proof. (5.11) gives one part of the equivalence and, since a connected Lie group satisfies the second axiom of countability, (5.15) gives the other.

It is not possible in (5.14) to drop the assumption that $\mathfrak{G}$ satisfies the second axiom of countability. For example, if G is a non-trivial connected Lie transformation group of a locally compact, locally connected space X, then in the discrete topology G would satisfy the hypotheses, but not the conclusion of (5.14). This shows that Theorem 9 of [1] is false as stated. The latter states a result similar to part of (5.14) in a special case. It is probably true when the second axiom of countabiilty is added as an assumption on the group G, however, the simple (but invalid) analytical proof given seems irreparable and topological arguments of the type we have used seem necessary.

6. Some dimension theory. In what follows a space stated to have dimension, finite or infinite, is assumed to have a separable metric topology. This is so that all the theorems of [3] will be valid. If X is a compact space and C a closed subset of X, then $H^q(X, C)$ will denote the q-dimensional Cêch cohomology group of X modulo C, and $\Pi^q(X, C)$ will denote

the q-dimensional cohomotopy classes of X relative to C, i.e., the homotopy classes of mappings of X into the q-sphere S^q which carry C into the north pole p_0. We wish first to show that if $\dim X = q > 0$, then these two sets are in one-to-one correspondence. Recall first that if $c \in X$ and $q > 0$, then $H^q(X, \{c\})$ is isomorphic to $H^q(X) = H^q(X, \emptyset)$ and that $\Pi^q(X, \{c\})$ is clearly in one-to-one correspondence with $\Pi^q(X) = \Pi^q(X, \emptyset)$. Now let $\tilde{X}$ be the space formed by identifying the closed subset C of the compact space X to a single point c, and let f be the natural projection of X on $\tilde{X}$. Then f is a relative homeomorphism of (X, C) on $(\tilde{X}, \{c\})$, so, by Theorem 5.4, pages 266 of [2], f^* is an isomorphism of $H^q(\tilde{X}, \{s\})$ with $H^q(X, C)$. It is also clear that $\Pi^q(X, \{C\})$ is in natural one-to-one correspondence with $\Pi^q(\tilde{X}, \{c\})$. Now suppose $\dim X = n < \infty$. Then $\dim(\tilde{X} - \{c\}) = \dim(X - C) \leqq n$ so, by Corollary 2, page 32 of [3], $\dim \tilde{X} \leqq n$. Now if $\dim \tilde{X} < n$, both $H^n(\tilde{X})$ and $\Pi^n(\tilde{X})$ contain just one point, while if $\dim \tilde{X} = n$, then $H^n(\tilde{X})$ and $\Pi^n(\tilde{X})$ are in one-to-one correspondence by Theorem VIII 2, page 149 of [3]. Suppose now that $n > 0$ and let us put all these one-to-one correspondences together:

$$H^n(X, C) \leftrightarrow H^n(\tilde{X}, \{c\}) \leftrightarrow H^n(\tilde{X}) \leftrightarrow \Pi^n(\tilde{X}) \leftrightarrow \Pi^n(\tilde{X}, \{c\}) \leftrightarrow \Pi^n(X, C).$$

6.1. LEMMA. *If X is a compact space of dimension n $(0 < n < \infty)$, then there is a one-to-one correspondence between $H^n(X, C)$ and $\Pi^n(X, C)$.*

6.2. THEOREM. *Let X be a compact space of finite dimension $n > 0$. Then there is a closed subset C of X for which there exists an essential map of the pair (X, C) into (S^n, p_0).*

Proof. Since the inessential maps of (X, C) into (S^n, p_o) are just those in the same homotopy class as the constant map $\bar{p}_o$, it suffices to show that $\Pi^n(X, C)$ contains more than one element for some closed subset C of X. In view of the lemma it suffices to find a closed subset C of X for which $H^n(X, C) \neq 0$. But if on the contrary $H^n(X, C) = 0$ for all closed subsets C of X, then it would follow from Theorem VIII 4, page 152 of [3] that $\dim X \leqq n - 1$, a contradiction.

6.3. BORSUK'S THEOREM. *Let Y be a closed subspace of a space X and C a closed subspace of Y. Let f and g be homotopic mappings of (Y, C) into (S^n, p_o). If f has an extension F over X relative to S^n, then there is an extension G of g over X such that F and G are homotopic mappings of (X, C) into (S^n, p_o).*

Proof. This is stated and proved as Theorem VI 5, page 86 of [3] for

the case $C=\emptyset$. However the proof given actually proves the more general relativized form stated above.

6.4. LEMMA. *Let Y be a closed subset of a space X and C a closed subset of Y. Let F be a map of (X,C) into (S^n,p_o) such that f, the restriction of F to Y, is an inessential map of (Y,C) into (S^n,p_o). Then there is a neighborhood V of Y such that the restriction of F to V is an inessential map of (V,C) into (S^n,p_o).*

Proof. Let g be the constant map $y\to p_o$ of (Y,C) into (S^n,p_o). By assumption f and g are homotopic mappings of (Y,C) into (S^n,p_o), hence by (6.3) there is an extension G of g over X such that F and G are in the same homotopy class as mappings of (X,C) into (S^n,p_o). *A fortiori* if V is any neighborhood of Y, then the restrictions of F and G to V are homotopic maps of (V,C) into (S^n,p_o), hence it suffices to find a neighborhood V of Y for which G restricted to V is an inessential map of (V,C) into (S^n,p_o). But clearly if U is any contractible neighborhood of p_o on S^n then $G^{-1}(U)=V$ works.

The following result, or at least closely related ones, are known. However the proof is short and we include it for completeness.

6.5. THEOREM. *Let C be a closed subspace of a compact space X and let f be an essential map of the pair (X,C) into (S^n,p_o). Then the family $\mathfrak{F}$ of closed subsets Y of X which include C and for which the restriction of f to Y is an essential map of (Y,C) into (S^n,p_o) contains a minimal element.*

Proof. By Zorn's lemma it suffices to show that the ordering of $\mathfrak{F}$ by inclusion is inductive, i.e., if Γ is a chain in $\mathfrak{F}$ and F is the intersection of the elements of Γ, then we must show that $F\in\mathfrak{F}$. Suppose on the contrary that $F\notin\mathfrak{F}$. Since clearly $C\subseteq F$ this means that f restricted to F is an inessential map of (F,C) into (S^n,p_o) and hence by (6.4) there is an open set V including F such that f restricted to V is an inessential map of (V,C) into (S^n,p_o). Now $\{Y-V\colon Y\in\Gamma\}$ is a chain of compact sets with empty intersection and hence $Y-V$ is empty for some $Y\in\Gamma$. But then $Y\subseteq V$ and hence the restriction of f to Y is an inessential map of (Y,C) into (S^n,p_o) so $Y\notin\mathfrak{F}$, contradicting $\Gamma\subseteq\mathfrak{F}$.

7. A criterion for Lie groups. This section contains the proof of a theorem reported by one of the authors in [4].

7.1. THEOREM. *Let G be a locally arcwise connected topological group. If X is a non-void compact, metrizable subspace of G of dimension $n < \infty$, then either X has an interior point or else there is an arc A in G such that AX is compact, metrizable and of dimension greater than n.*

Proof. We note first that if A is any arc in G, then AX is compact and metrizable. In fact, A is the continuous image of the unit interval I under some map σ so that AX is the continuous image of the compact, metrizable space $I \times X$ under the continuous map $(t,x) \to \sigma(t)x$. The desired result now follows from Satz IX, § 3, Chapter II of [7].

If $n = 0$ then either G is discrete, in which case every point of X is an interior point, or else G has a non-trivial arc A in which case AX has a non-trivial arc and therefore is of dimension greater than or equal to one.

Now suppose $n > 0$. By (6.2) we can find a closed subset C of X for which there exists an essential map f of the pair (X, C) into (S^n, p_o). By (6.5) there is a closed subset X' of X including C such that f restricted to X' is an essential map of the pair (X', C) into (S^n, p_o), but for any non-void open subset U of X' disjoint from C the restriction of f to $X' - U$ is an inessential map of $(X' - U, C)$ onto (S^n, p_o). Without loss of generality we can assume that $e \in X' - C$, for in any case this can be arranged by a translation.

Let V be an arcwise connected neighborhood of e such that $V^{-1}\bar{V}$ is disjoint from C. Suppose now that X has no interior points. Then certainly V^{-1} is not included in X', so we can find a continuous map σ of the unit interval into G such that $\sigma(0) = e$, $A = \text{range of } \sigma \subseteq V$, and $\sigma(1)^{-1} \notin X'$. Then $e \notin \sigma(1)X'$, hence since X' is compact, we can find an open neighborhood U of e with $U \subseteq V$ and $\bar{U}$ disjoint from $\sigma(1)X'$. By the choice of X' the restriction of f to $X' - U$ is homotopic, as a map of the pair $(X' - U, C)$ into (S^n, p_o), to the constant mapping $x \to p_o$. By (6.3) it follows that there is a map g of the pair (X', C) into (S^n, p_o) such that $g(x) = p_o$ for $x \in X' - U$ which is homotopic to the restriction of f to X' and therefore is essential.

Now note that $AC \cup \sigma(1)X'$ is disjoint from $\bar{U}$. In fact, $\bar{U}$ was chosen disjoint from $\sigma(1)X'$ while, since $A \subseteq V$ and $V^{-1}\bar{V}$ is disjoint from C, it follows that AC is disjoint from $\bar{V}$ and *a fortiori* from $\bar{U}$. Then since $g(x) = p_o$ for $x \in X' - U$, it follows that the defining conditions $h(x) = p_o$ for $x \in AC \cup \sigma(1)X'$, $h(x) = g(x)$ for $x \in X'$ are non-contradictory and define a continuous mapping h of $X' \cup AC \cup \sigma(1)X'$ into S^n. It follows from the essentiality of g that h does not have a continuous extension over

AX' relative to S^n. In fact, if $\bar{h}$ were such an extension of h, then $H: I \times X' \to S^n$, defined by $H(t,x) = \bar{h}(\sigma(t)x)$ would be a homotopy of g (considered as a map of the pair (X', C) into (S^n, p_0)) with the constant map $x \to p_0$. It follows *a fortiori* that h does not have a continuous extension over AX relative to S^n. Since as we have already seen AX is metrizable, we can apply the theorems of [3]. In particular, by Theorem VI 4, page 83 of [3], the existence of a continuous mapping of a closed subspace of AX into S^n which does not admit a continuous extension over all of AX implies that $\dim(AX) > n$.

7.2. THEOREM. *A locally arcwise connected topological group G in which the compact metrizable subspaces are of bounded dimension is a Lie group.*

Proof. Let n be the least upper bound of the dimensions of the compact, metrizable subspaces of G and let X be a compact metrizable subspace of G of dimension n. By (7.1) X has an interior point g. Then $V = g^{-1}X$ is a compact n-dimensional neighborhood of the identity in G. Hence G is a locally connected, locally compact, n-dimensional topological group and, by the theorem on page 185 of [5], G is a Lie group.

7.3. COROLLARY. *If G is a topological group in which the compact, metrizable subspaces are of bounded dimension, then the associated locally arcwise connected group of G (Definition 3.3) is a Lie group.*

Proof. Since the topology of G^*, the associated locally arcwise connected group of G, is stronger than the topology of G, the compact metrizable subspaces of G^* are also compact metrizable subspaces of G and hence have bounded dimension.

7.4. COROLLARY. *Let G be a topological group in which some neighborhood of the identity admits a continuous one-to-one map into a finite dimensional metric space. Then the associated locally arcwise connected group of G is a Lie group and in particular if G is locally arcwise connected, then G is a Lie group.*

Proof. Let f be a continuous one-to-one map of a closed neighborhood V of the identity into a finite dimensional metric space X. Given a compact subspace K of G and $k \in K$, $kV \cap K$ is a compact neighborhood of k in K. The map $h \to f(k^{-1}h)$ maps $kV \cap K$ homeomorphically onto a compact subspace of X. It follows that $kV \cap K$ is of dimension less than or equal to n. Since each point of K has a neighborhood of dimension less than or equal to n, it follows that $\dim(K) \leq n$ and (7.4) now follows from (7.3).

8. A criterion for Lie transformation groups.

8.1. *Definition.* A set Φ of homeomorphisms of a space X will be said to be *faithfully represented* by a subset F of X if $\phi \in \Phi$ and $\phi(x) = x$ for all $x \in F$ implies ϕ is the identity map of X.

8.2. THEOREM. *Let $\mathfrak{G} = (G, \mathfrak{T})$ be a Lie group whose underlying group G is a group of homeomorphism of a space X and whose topology $\mathfrak{T}$ is admissibly strong. Then some neighborhood of the identity in $\mathfrak{G}$ is faithfully repersented by a finite subset F of X. In fact, if $\dim \mathfrak{G} = n$ then F can be taken to have n or fewer points.*

Proof. Given a finite subset F of X, let

$$G_F = \{g \in G : g(x) = x \text{ for all } x \in F\}.$$

Each G_F is a closed subgroup of $\mathfrak{G}$ and hence a Lie group. It will suffice to prove that for some F containing n or fewer points, G_F is zero-dimensional and hence discrete. We prove this by showing that if $\dim G_F > 0$, then for some $x \in X$ we have $\dim G_{F \cup \{x\}} < \dim G_F$. In fact, let $\phi : t \rightarrow \phi_t$ be a non-trivial one-parameter subgroup of G_F and let t be a real number such that $\phi_t \neq e$. Choose $x \in X$ such that $\phi_t(x) \neq x$. Then $G_{F \cup \{x\}}$ is a subgroup of G_F and ϕ is a one-parameter subgroup of G_F but not of $G_{F \cup \{x\}}$.

There is a partial converse to (8.2), namely,

8.3. THEOREM. *Let $\mathfrak{G} = (G, \mathfrak{T})$ be a topological group whose underlying group G is a group of homeomorphisms of a finite dimensional metric space X and whose topology $\mathfrak{T}$ is admissibly strong. If some neighborhood of the identity in $\mathfrak{G}$ is faithfully represented by a finite subset of X, then the associated locally arcwise connected group of $\mathfrak{G}$ is a Lie group, and in particular if $\mathfrak{G}$ is locally arcwise connected, it is a Lie group.*

Proof. Let U be a neighborhood of the identity in $\mathfrak{G}$ faithfully represented by a finite subset $x_1, \cdots, x_n$ of X and choose a neighborhood V of the identity with $V^{-1}V \subseteq U$. Then $f : g \rightarrow (g(x_1), \cdots, g(x_n))$ is a continuous map of V into X^n. Moreover, if $f(g) = f(h)$ for $g, h \in V$ then $g^{-1}h(x_i) = g^{-1}g(x_i) = x_i$, $i = 1, 2, \cdots, n$; since $g^{-1}h \in U$, it follows that $g^{-1}h = e$ or $g = h$. Thus f is one-to-one and since $\dim X^n \leqq n \dim X < \infty$, the proposition follows from (7.4).

8.4. THEOREM. *Let X be a locally compact, locally connected, finite dimensional metric space. A necessary and sufficient condition for a group G of homeomorphisms of X to be a Lie transformation group of X with the*

modified compact-open topology as its Lie topology is that some modified compact-open neighborhood of the identity in G be faithfully represented by a finite subset of X.

Proof. Necessity follows from (8.2). Since the modified compact-open topology for G is admissibly strong (5.10) and makes G a locally arcwise connected topological group (5.12), sufficiency follows from (8.3).

Since the modified compact-open neighborhoods of the identity are generally impossible to determine while the compact-open neighborhoods of the identity are explicitly given, the following corollary to (8.4) is a more useful criterion than the theorem itself.

8.5. COROLLARY. *Let G be a group of homeomorphisms of a locally compact, locally connected, finite dimensional metric space X. If some compact-open neighborhood of the identity in G is faithfully represented by a finite subset of X, then G is a Lie transformation group of X and the modified compact-open topology for G is its Lie topology. In particular, if G itself is faithfully represented by a finite subset of X, it is a Lie transformation group of X with the modified compact-open topology as its Lie topology.*

Proof. A compact-open neighborhood of the identity in G is *a fortiori* a modified compact-open neighborhood of the identity.

It is perhaps in order here to remark on the relevance (or rather irrelevance) of the various metrizability assumptions we have made in Section 6 and thereafter. In general these have been made to justify the use made of theorems proved in [3], where all spaces are taken to be separable metric. However, since all the spaces to which dimension arguments are directly applied are, in this paper, compact, it is possible as the referee suggests to use the Lebesque definition of dimension in terms of the minimum number of interesecting sets in small open coverings, and drop all references to metrizability. For the results in Section 6 and for Theorem 7.1 this would strengthen our results. However for the main theorem, Theorem 7.2, dropping the reference to metrizability would at least formally weaken the theorem. Using the referee's suggestion, Corollary 7.4 can be strengthened by replacing metric by compact in its statement. That this gives a really stronger result follows from the fact that a finite dimensional metric space can be imbedded in a finite dimensional compact space. In fact the referee notes that locally compact can replace metric in Corollary 7.4, provided the dimension of a locally compact space is defined as the supremum of the

dimensions of its compact subspaces. This is immediate from the proof of this corollary. Having noted this last fact it follows that in Theorems 8.3 and 8.4 and in Corollary 8.5 we can drop the assumption that X is metric.

9. A conjecture. Let X be a connected manifold satisfying the second axiom of countability. Let G be a connected Lie transformation group of X, or what is the same (5.16), a group of homeomorphisms of X that is a connected Lie group in its modified compact-open topology. Let $\bar{G}$ be the closure of G, relative to the compact-open topology, in the group of all homeomorphisms of X. Then we conjecture that $\bar{G}$ is also a Lie transformation group of X and that the Lie topology of $\bar{G}$ is its compact-open topology.

If this is so, then the structure of the class of connected Lie transformation groups of X is very clear. On the one hand, there are those groups of homeomorphisms of X which are connected Lie groups in their compact-open topology, and all others are analytic subgroups of these. The validity of this structure theorem would have many interesting consequences.

If M is a differentiable manifold with X as its underlying topological manifold, then we define a Lie transformation group of M to be a Lie transformation group of X consisting entirely of differentiable homeomorphisms. The differentiable structure of M allows one to develop an infinitesimal characterization of Lie transformation groups of M in terms of vector fields on M. This is carried out by one of the authors in a recent Memoir of the American Mathematical Society [6].

HARVARD UNIVERSITY.

REFERENCES.

[1] R. F. Arens, "Topologies for homeomorphism groups," *American Journal of Mathematics*, vol. 68 (1946), pp. 593-610.

[2] S. Eilenberg and N. Steenrod, *Foundations of Algebraic Topology*, Princeton, 1952.

[3] W. Hurewicz and H. Wallman, *Dimension Theory*, Princeton, 1948.

[4] A. M. Gleason, "Arcs in locally compact groups," *Proceedings of the National Academy of Sciences*, vol. 38 (1950), pp. 663-667.

[5] D. Montgomery and L. Zippin, *Topological Transformation Groups*, Interscience, New York, 1955.

[6] R. S. Palais, *A Global Formulation of the Lie Theory of Transformation Groups*, Memoirs of the American Mathematical Society 22, 1957.

[7] P. Alexandroff and H. Hopf, *Topologie*, I, Berlin, Springer, 1935.

[8] N. Bourbaki, *Topology Générale*, Chapt. IX, Hermann, Paris, 1948.

ON THE BIRATIONAL EQUIVALENCE OF CURVES UNDER SPECIALIZATION.*

By Wei-Liang Chow[1] and Serge Lang.

Let k be a field with a discrete valuation, let p be the maximal prime ideal in the valuation ring on k, and let $\bar{k}$ be the residue field of k. Let S^n be the projective space of dimension n in the algebraic geometry over a universal domain containing k, and denote by $\bar{S}^n$ the projective space of dimension n over a universal domain containing $\bar{k}$; if Z is any cycle in S^n, rational over k, we shall denote by $\bar{Z}$ the cycle obtained from Z by reduction modulo p, in the sense of Shimura [4], and we shall call $\bar{Z}$ the specialization of Z (with respect to the given valuation of k). This definition applies also in case Z is a variety, which is then to be considered as a prime cycle; the the specialization $\bar{Z}$ is then a positive cycle, but in general not necessarily a variety. According to the generalization of the Principle of Degeneration recently proved by Chow, the specialization of a variety is always connected, but we shall not need this fact here.

Consider now two non-singular curves C_1 and C_2 in S^n, both defined over k, and assume that they are birationally equivalent over k. A theorem of Deuring ([2], Satz 3) can be expressed geometrically by saying that if C_1 and C_2 have genus 1 and if the cycles $\bar{C}_1$ and $\bar{C}_2$ are also non-singular curves of genus 1, then $\bar{C}_1$ and $\bar{C}_2$ are birationally equivalent $\bar{k}$. In this note we shall generalize this result to non-singular curves of arbitrary genus and also to abelian varieties; in fact, we shall first prove the result for abelian varieties, and then deduce the result for curves by embedding them in their Jacobians. Our method is quite different from Deuring's, and is based on the "compatibility" of the Chow construction of the Jacobian and the canonical mapping.

Let A be an abelian variety in S^n, defined over k. We shall say that the specialization $\bar{A}$ of A is *non-degenerate* if $\bar{A}$ is an abelian variety defined over $\bar{k}$ and if the law of composition in $\bar{A}$ is the specialization of the law of

---

* Received January 24, 1957.

[1] Work of W. L. Chow was supported in part by a National Science Foundation Grant, NSF-3022.

composition in A. Specifically, this means the following. Let G be the graph of the law of composition in A, so that G is a subvariety in $A \times A \times A$; then our condition can be expressed by saying that both $\bar{A}$ and $\bar{G}$ are varieties (in $\bar{S}^n$ and $\bar{S}^n \times \bar{S}^n \times \bar{S}^n$ respectively) and that $\bar{G}$ defines an abelian law of composition in $\bar{A}$. We observe that if e is the unit element in A, then $\bar{e}$ is the unit element in $\bar{A}$; in fact, e is characterized by the condition that $\mathrm{pr}_{32}(e \times A \times A))$ is contained in the diagonal in $A \times A$, and this condition is evidently preserved under the specialization.

THEOREM 1. *Let A and B be abelian varieties in S^n, and let T be a birational isomorphism between A and B; we assume that A, B, and T are all defined over k. If the specializations $\bar{A}$ and $\bar{B}$ are non-degenerate, then the specialization $\bar{T}$ is a variety and defines a birational isomorphism between $\bar{A}$ and $\bar{B}$.*

Proof. It is clear that $\bar{A} \times \bar{B}$ is the specialization of $A \times B$ and is non-degenerate. If we denote the graph of T by the same symbol, then the support $|\bar{T}|$ of $\bar{T}$ is an algebraic subgroup in $\bar{A} \times \bar{B}$; this follows from the fact that T is an abelian subvariety in $A \times B$, and the fact that the law of composition in $\bar{A} \times \bar{B}$ is the specialization of the law of composition in $A \times B$. Let $\bar{T}_0$ be a component in $\bar{T}$ which contains the unit element; then $\bar{T}_0$ is an abelian subvariety in $\bar{A} \times \bar{B}$, and every component in $|\bar{T}|$ and hence also every component in $\bar{T}$ is obtained from $\bar{T}_0$ by a translation. By assumption, we have $\mathrm{pr}_1 T = A$ and $\mathrm{pr}_2 T = B$; hence, by specialization, we have $\mathrm{pr}_1 \bar{T} = \bar{A}$ and $\mathrm{pr}_2 \bar{T} = \bar{B}$. Since all cycles obtained from $\bar{T}_0$ by translations must have the same projections on $\bar{A}$ and $\bar{B}$, we conclude readily that $\mathrm{pr}_1 \bar{T}_0 = \bar{A}$ and $\mathrm{pr}_2 \bar{T} = \bar{B}$, and that $\bar{T}_0$ is the only component in $\bar{T}$ and has the coefficient 1. This shows that $\bar{T} = \bar{T}_0$ defines a birational transformation between $\bar{A}$ and $\bar{B}$, which, since it preserves the unit elements, must be an isomorphism.

We turn now to the case of the curves C_1 and C_2 considered at the beginning of this note. Let C be a non-singular curve of genus $g(C)$ in S^n, defined over k; we shall say that the specialization $\bar{C}$ is *non-degenerate* if $\bar{C}$ is a non-singular curve and hence has the same genus $g(C)$ as C. We assume that there exists a birational correspondence F between C_1 and C_2, defined over k, and that the specialization $\bar{C}_1$ and $\bar{C}_2$ are both non-degenerate. Since the conditions $\mathrm{pr}_1 F = C_1$ and $\mathrm{pr}_2 F = C_2$ imply by specialization the conditions $\mathrm{pr}_1 \bar{F} = \bar{C}_1$ and $\mathrm{pr}_2 \bar{F} = \bar{C}_2$, we conclude that $\bar{F}$ is either a birational correspondence between $\bar{C}_1$ and $\bar{C}_2$, or has the form $x \times \bar{C}_2 + \bar{C}_1 \times y$, for some points x and y in $\bar{C}_1$ and $\bar{C}_2$ respectively. In case $g(C_1) = g(C_2) = 0$, it can be easily

seen by examples that both possibilities can occur; however, in this case the assertion of Deuring is trivially true. In the next theorem we shall show that, except in this trivial case, the second possibility cannot occur; this implies that the specialization $\bar{F}$ is itself already such a birational correspondence.

THEOREM 2. *Let C_1 and C_2 be non-singular curves of genus > 0 in S^n, and let F be a birational correspondence between C_1 and C_2; we assume that C_1, C_2, and F are all defined over k. If the specializations $\bar{C}_1$ and $\bar{C}_2$ are non-degenerate, then the specialization $\bar{F}$ is a birational correspondance between $\bar{C}_1$ and $\bar{C}_2$.*

Proof. The theorem being "geometric," we may assume that the curves C_1 and C_2 have rational points over k, and hence the curves $\bar{C}_1$ and $\bar{C}_2$ have rational points over $\bar{k}$. In fact, we can enlarge k by a finite algebraic extension if necessary, and extend the discrete valuation in k to a discrete valuation of this extension; it is easily seen that the validity of our theorem over this extension will imply its validity in the original form.

For $i=1,2$, we consider the Jacobian J_i of C_i and the canonical mapping f_i of C_i into J_i with reference to a rational point p_i in C_i (i.e., $f_i(p_i)=e_i$, e_i being the unit element in J_i), both J_i and f_i being constructed as in Chow [1]. It is essential for our purpose that J_i and f_i are constructed in this particular way, which we shall call the Chow construction, not just some arbitrarily chosen projective models of J_i and the corresponding mapping f_i. The reason lies in the fact that the Chow construction is "compatible" with specializations, at least when they are non-degenerate. Specifically, this means that if the specialization $\bar{C}_i$ is non-degenerate, then the specialization $\bar{J}_i$ (which is a cycle in $\bar{S}^{t_i}$ if S^{t_i} is the ambient projective space of J_i) of the abelian variety J_i is also non-degenerate, and $\bar{J}_i$ and $\bar{f}_i$ are respectively the Jacobian and the canonical mapping of $\bar{C}_i$. This fact is a special case of a more general "compatibility" theorem of Igusa [3] (Igusa treats only the algebro-geometric specializations, but it is easily seen that his proof is also valid for the more general specializations considered here at least in the non-degenerate case). For the sake of convenience, we shall assume that $p_2=F(p_1)$; then there exists a birational isomorphism T between J_1 and J_2, defined over k, such that $T\circ f_1=f_2\circ F$. If we denote by f the mapping $f_1\times f_2$ of $C_1\times C_2$ into $J_1\times J_2$, and if we denote the graphs of F and T by the same symbols, then the above relation implies that $f(F)$ is contained in T. Since the specialization $\bar{f}=\bar{f}_1\times\bar{f}_2$ is a rational mapping of $\bar{C}_1\times\bar{C}_2$ into $\bar{J}_1\times\bar{J}_2$, it follows that the cycle $\bar{f}(\bar{F})$ is defined and is contained in $|\bar{T}|$.

By Theorem 1, $\bar{T}$ is an abelian subvariety in $\bar{J}_1 \times \bar{J}_2$ (it is in fact sufficient to know that $|\bar{T}|$ is an algebraic group in $\bar{J}_1 \times \bar{J}_2$) and is a proper subset in $\bar{J}_1 \times \bar{J}_2$. If $\bar{F}$ is not a birational correspondence between $\bar{C}_1$ and $\bar{C}_2$, then by a previous remark we must have $\bar{F} = \bar{p}_1 \times \bar{C}_2 + \bar{C}_1 \times \bar{p}_2$ (we observe that $|\bar{F}|$ must contain the point $\bar{p}_1 \times \bar{p}_2$) and hence

$$\bar{f}(\bar{F}) = \bar{e}_1 \times \bar{f}_2(\bar{C}_2) + \bar{f}_1(\bar{C}_1) \times \bar{e}_2.$$

This shows that $|\bar{f}(\bar{F})|$ generates the entire variety $\bar{J}_1 \times \bar{J}_2$, in contradiction to the fact that $\bar{f}(\bar{F})$ is contained in $\bar{T}$. This proves our theorem.

THE JOHNS HOPKINS UNIVERSITY,
COLUMBIA UNIVERSITY.

REFERENCES.

[1] W. L. Chow, "The Jacobian variety of an algebraic curve," *American Journal of Mathematics*, vol. 76 (1954), pp. 453-476.

[2] M. Deuring, "Die Zetafunktion einer algebraischen Kurve vom Geschlechte Eins (II)," *Nachrichten der Akademie der Wissenschaften in Göttingen*, 1955, pp. 13-42.

[3] J. Igusa, "Fibre systems of Jacobian varieties," *American Journal of Mathematics*, vol. 78 (1956), pp. 171-199.

[4] G. Shimura, "Reduction of algebraic varieties with respect to a discrete valuation of the basic field," *ibid.*, vol. 77 (1955), pp. 131-176.

FOURIER TRANSFORMS ON A SEMISIMPLE LIE ALGEBRA II.*

By Harish-Chandra.

1. Introduction. The purpose of this paper is to complete the proofs of two results (Theorems 2 and 3) which have already been mentioned in [4(e), § 1]. For the first one, the problem reduces to showing that the constants of the Corollary to Lemma 29 of [4(e)] are all equal. It is obvious that this can be done only by taking the singular elements of $\mathfrak{g}_0$ into account. However, as we shall see, it is enough to consider only the best among these, namely, the semiregular ones (see Section 2). Let $\mathfrak{g}_0'$ be the set of all regular elements in $\mathfrak{g}_0$. We first show that a given semiregular element H_0 of $\mathfrak{g}_0$ can lie in the closure of at most three distinct connected components of $\mathfrak{g}_0'$. Let $\mathfrak{l}_0$ be the real semisimple Lie algebra of dimension 3 spanned by the elements H, X, Y satisfying the relations $[H,X]=2X$, $[H,Y]=-2Y$, $[X,Y]=H$. Then zero is the only semiregular element in $\mathfrak{l}_0$ and certain computations on $\mathfrak{g}_0$ around H_0 can be reduced to similar computations on $\mathfrak{l}_0$ around zero (see Lemma 8 and its Corollary). Now the distribution T' of [4(e), § 7] satisfies certain differential equations [4(e), Lemma 30]. If we transcribe these equations on $\mathfrak{l}_0$, they become simple enough to be handled directly (see Lemma 12). In this way, one proves that T' coincides with a constant on some neighborhood of H_0. Theorem 2 now follows from the fact that the union of regular and semiregular elements of $\mathfrak{g}_0$ is connected (Lemma 13). Once Theorem 2 is proved, Theorem 3 can be obtained without much difficulty from Lemma 41 of [4(e)].

For the convenience of the reader all calculations on $\mathfrak{l}_0$ have been collected together in the Appendix (Section 6).

The main results of this paper have been announced in a short note [4(f)].

2. Transformation of certain integrals. We keep to the notation of [4(d), (e)]. Call an element $X \in \mathfrak{g}$ *semiregular* if (1) $\operatorname{ad} X$ is semisimple and (2) the centralizer of X in $\mathfrak{g}$ is of dimension $l+2$ (where $l = \operatorname{rank} \mathfrak{g}$). Let $\mathfrak{h}_0$ be a Cartan subalgebra of $\mathfrak{g}_0$ such that $\theta(\mathfrak{h}_0)=\mathfrak{h}_0$ and let H_0 be a semiregular element in $\mathfrak{h}_0$. Define P, P_+, P_0 and P_- as in Section 5 of [4(e)].

* Received July 30, 1956.

Then since H_0 is semiregular, it is obvious that $\alpha_0(H_0)=0$ for exactly one root $\alpha_0 \in P$, and the centralizer of H_0 in $\mathfrak{g}$ is $\mathfrak{z}=\mathfrak{h}+CX_{\alpha_0}+CX_{-\alpha_0}$. Let Q be the set of those roots $\alpha \in P$ for which α^2 takes only real values on $\mathfrak{h}_0$. We claim $\alpha_0 \in Q$. For otherwise, suppose $\alpha_0 \notin Q$ and put $\beta=-\theta\alpha_0$. Then β is also a root and $\beta(H)=\operatorname{conj}\alpha_0(H)$ $(H\in\mathfrak{h}_0)$. This implies that $\beta(H_0)=0$, and therefore $X_\beta \in \mathfrak{z}$. On the other hand $\beta \neq \pm\alpha_0$ since $\alpha_0 \notin Q$. Therefore it is obvious that $X_\beta \notin \mathfrak{z}$, and we get a contradiction. Hence $\alpha_0 \in Q$. Moreover, it is clear that $\tilde{\theta}(\mathfrak{z})=\mathfrak{z}$, and therefore (see [4(e), Lemma 10]) $\mathfrak{z}$ is reductive in $\mathfrak{g}$ and $\mathfrak{z}=\sigma+\mathfrak{l}$, where σ is the center and $\mathfrak{l}=[\mathfrak{z},\mathfrak{z}]$ $=CH_{\alpha_0}+CX_{\alpha_0}+CX_{-\alpha_0}$ the derived algebra of $\mathfrak{z}$. Obviously σ is the set of all $H\in\mathfrak{h}$, where $\alpha_0(H)=0$. Put $\sigma_0=\sigma\cap\mathfrak{h}_0$. Then $\mathfrak{z}_0=\mathfrak{z}\cap\mathfrak{g}_0=\sigma_0+\mathfrak{l}_0$, where $\mathfrak{l}_0=\mathfrak{l}\cap\mathfrak{g}_0$. Now $\dim_R \mathfrak{z}_0=\dim\mathfrak{z}=l+2$ and $\dim_R\sigma_0=l-1$. Hence $\dim_R\mathfrak{l}_0=3$, and this shows that $\mathfrak{l}_0$ is a real form of the complex semisimple Lie algebra $\mathfrak{l}$. Put $H_{\alpha_0}'=2\alpha_0(H_{\alpha_0})^{-1}H_{\alpha_0}$ and define Q_+ as in [4(e), § 12]. We select X_{α_0} and $X_{-\alpha_0}$ in accordance with Corollary 1 of Lemma 46 of [4(e)]. Then $\mathfrak{l}_0=RH_{\alpha_0}'+RX_{\alpha_0}+RX_{-\alpha_0}$ if $\alpha_0\in Q_+$ and

$$\mathfrak{l}_0=R(-1)^{\frac{1}{2}}H_{\alpha_0}'+R(X_{\alpha_0}-X_{-\alpha_0})+R(-1)^{\frac{1}{2}}(X_{\alpha_0}+X_{-\alpha_0}) \text{ if } \alpha_0\in P_-.$$

In the latter case, $\mathfrak{l}_0$ is compact (see Section 6). Finally,

$$\mathfrak{l}_0=R(-1)^{\frac{1}{2}}H_{\alpha_0}'+R(-1)^{\frac{1}{2}}(X_{\alpha_0}-X_{-\alpha_0})+R(X_{\alpha_0}+X_{-\alpha_0}) \text{ if } \alpha_0\in P_0.$$

Let $\mathfrak{l}'$ be the set of those points in $\mathfrak{l}$ which are regular (with respect to $\mathfrak{l}$). Put $\mathfrak{l}_0'=\mathfrak{l}'\cap\mathfrak{l}_0$.

Let $\mathfrak{q}$ be the orthogonal complement of $\mathfrak{z}$ in $\mathfrak{g}$. Then $\mathfrak{q}=\sum_{\substack{\alpha\in P\\ \alpha\neq\alpha_0}}(CX_\alpha+CX_{-\alpha})$ and $[\mathfrak{z},\mathfrak{q}]\subset\mathfrak{q}$. Let $\zeta(Z)$ $(Z\in\mathfrak{z})$ denote the determinant of the restriction of $\operatorname{ad}Z$ on $\mathfrak{q}$. Then ζ is a polynomial function on $\mathfrak{z}$ which takes only real values on $\mathfrak{z}_0$ and $\zeta(H)=(-1)^{r-1}\prod_{\substack{\alpha\in P\\ \alpha\neq\alpha_0}}\alpha(H)^2$ $(H\in\mathfrak{h})$, where r is the number of roots in P. This shows that $\zeta(H_0)\neq 0$. Let $\mathfrak{z}_0'$ be the set of those points $Z\in\mathfrak{z}_0$ where $\zeta(Z)\neq 0$. Select an open connected neighborhood U of H_0 in $\mathfrak{z}_0'$ and consider the factor space $G^*=G/\Xi$, where G is the (connected) adjoint group of $\mathfrak{g}_0$ and Ξ is the centralizer of H_0 in G. We denote by $x\to x^*$ $(x\in G)$ the natural mapping G on G^*. Choose an open connected neighborhood V^* of 1^* in G^*. Then if V^* is sufficiently small, we can define (see Chevalley [2, p. 111]) an analytic mapping ϕ of V^* into G with the following properties: (1) $(\phi(x^*))^*=x^*$ $(x^*\in V^*)$ and (2) ϕ is regular on V^*. Put $V=\phi(V^*)$ and consider the mapping ψ: $(x^*,Z)\to\phi(x^*)Z$ $(x^*\in V^*, Z\in U)$ of $V^*\times U$ into $\mathfrak{g}_0$. We shall prove that ψ is regular. For

$Y \in \mathfrak{g}_0$ and $X^* \in V^*$, put[1] $Y_{x^*}{}^* = d\tau_x(Y)$, where $x = \phi(x^*)$ and τ is the natural mapping of G on G^*. Identify the tangent space of $V^* \times U$ at (x^*, Z) with $\mathfrak{g}_0$ under the linear isomorphism

$$Y + X \leftrightarrow Y_{x^*}{}^* \times X \qquad (Y \in \mathfrak{q}_0, X \in \mathfrak{z}_0).$$

Then a simple calculation shows that

$$d\psi_{x^*,Z}(Y + X) = \phi(x^*)X - \phi(x^*) \operatorname{ad} Z(d\phi_{x^*} Y_{x^*}{}^*).$$

But since $\tau\phi$ is the identity mapping on V^*, $d\tau_x(d\phi_{x^*} Y_{x^*}{}^* - Y) = 0$ if $x = \phi(x^*)$, and therefore $d\phi_{x^*} Y_{x^*}{}^* \equiv Y \bmod \mathfrak{z}_0$. Hence if $D_{x^*,Z} = (\phi(x^*))^{-1} d\psi_{x^*,Z}$, it is clear that

$$D_{x^*,Z}X = X \quad (X \in \mathfrak{z}_0), \qquad D_{x^*,Z}Y \equiv -(\operatorname{ad} Z)Y \quad (Y \in \mathfrak{q}_0),$$

and so we conclude that $|\det(d\psi_{x^*,Z})| = |\det D_{x^*,Z}| = |\zeta(Z)| \neq 0$, since $U \subset \mathfrak{z}_0'$. This proves that ψ is everywhere regular on $V^* \times U$. Therefore, since the dimensions of the two manifolds $V^* \times U$ and $\mathfrak{g}_0$ are equal, it follows that ψ is also open and univalent provided V^* and U are sufficiently small. Put $N = \psi(V^* \times U)$. Then N is an open connected neighborhood of H_0 in $\mathfrak{g}_0$ and ψ defines an analytic isomorphism of $V^* \times U$ with N.

Since $\mathfrak{z}_0$ is reductive, the group Ξ is unimodular, and therefore (see Weil [8, p. 45]) there exists an invariant measure dx^* on G^*. Let dX and dZ denote the Euclidean measures on $\mathfrak{g}_0$ and $\mathfrak{z}_0$, respectively. Then the above calculation shows that dx^* can be so normalized that

$$\int_N f(X)\,dX = \int_{V^* \times U} f(\psi(x^*, Z))\,|\zeta(Z)|\,dx^* dZ \qquad (f \in C_c(N)).$$

Select once for all an element $a_0 \in C_c^\infty(V^*)$ such that $\int a_0(x^*)\,dx^* = 1$, and for any $\gamma \in C_c^\infty(U)$ define $f_\gamma \in C_c^\infty(N)$ by $f_\gamma(\psi(x^*, Z)) = a_0(x^*)\gamma(Z)$ $(x^* \in V^*, Z \in U)$. Then the following result is obvious.

LEMMA 1. *Let $'U$ be any open subset of U and put $'N = \psi(V^* \times {}'U)$. Then*

$$\int_{'N} f_\gamma(X)\,dX = \int_{'U} \gamma\,|\zeta|\,dZ \text{ for all } \gamma \in C_c^\infty('U).$$

On the other hand, it is clear that σ_0 and $\mathfrak{l}_0$ are mutually orthogonal. Let U_σ and $U_{\mathfrak{l}}$, respectively, denote the orthogonal projections of U in σ_0 and $\mathfrak{l}_0$. Then by replacing U, if necessary, by a smaller open connected

[1] We follow here the terminology of Chevalley [2]. The present computation is very similar to the one performed during the proof of Lemma 39 of [4(c)].

neighborhood of H_0 in $\mathfrak{z}_0'$, we can assume that (1) $U = U_\sigma + U_{\mathfrak{l}}$, (2) U_σ is convex and $\alpha(H) \neq 0$ for any $H \in U_\sigma$ and $\alpha \neq \alpha_0$ in P, (3) if $Z \in U_{\mathfrak{l}}$ then tZ $(0 \leqq t \leqq 1)$ is also in $U_{\mathfrak{l}}$. Define the polynomial η as in [4(e), § 6] and let $\mathfrak{g}_0'$ be the set of all regular elements in $\mathfrak{g}_0$. Put $U' = U \cap \mathfrak{g}_0'$ and $U_{\mathfrak{l}}' = U_{\mathfrak{l}} \cap \mathfrak{l}_0'$. Then we claim that $U' = U_\sigma + U_{\mathfrak{l}}'$. This is seen as follows. Let λ be an indeterminate and I the identity maping of $\mathfrak{g}$. Then it is obvious that

$$\det(\lambda I - \operatorname{ad} Z) = \det(\lambda I - \operatorname{ad} Z)_{\mathfrak{q}} \det(\lambda I - \operatorname{ad} Z)_{\mathfrak{z}} \qquad (Z \in \mathfrak{z}),$$

where the subscripts signify restrictions on the corresponding subspaces. Now suppose $Z \in U$ and $Z = H + Y$, where $H \in U_\sigma$ and $Y \in U_{\mathfrak{l}}$. Then it is evident that

$$\det(\lambda I - \operatorname{ad} Z)_{\mathfrak{z}} = \lambda^{l-1} \det(\lambda I - \operatorname{ad} Y)_{\mathfrak{l}},$$

and therefore $\det(\lambda I - \operatorname{ad} Z) = \lambda^{l-1} \det(\lambda I - \operatorname{ad} Z)_{\mathfrak{q}} \det(\lambda I - \operatorname{ad} Y)_{\mathfrak{l}}$. The coefficient of λ^l on the left side is $(-1)^r \eta(Z)$, where r is the number of roots in P. On the other hand, the constant term in $\det(\lambda I - \operatorname{ad} Z)_{\mathfrak{q}}$ is $(-1)^q \zeta(Z)$, where $q = \dim \mathfrak{q}$. Moreover, since $\mathfrak{l}$ is of rank 1, λ divides $\det(\lambda I - \operatorname{ad} Y)_{\mathfrak{l}}$, and the coefficient of λ in it is $-\frac{1}{2}\Omega(Y)$, where Ω is the Casimir polynomial of $\mathfrak{l}$ (see Section 6). This means that

$$(-1)^r \eta(Z) = \tfrac{1}{2}(-1)^{q+1} \zeta(Z) \Omega(Y).$$

Therefore, since ζ is never zero on U, $Z \in U'$ if and only if $Y \in \mathfrak{l}_0'$, and this proves that $U' = U_\sigma + U_{\mathfrak{l}}'$. In view of the above condition imposed on $U_{\mathfrak{l}}$, it follows (see Appendix, § 6) that $U_{\mathfrak{l}}'$ is connected or it has exactly three connected components according as $\mathfrak{l}_0$ is compact or not. Put $N' = N \cap \mathfrak{g}_0'$. Then $N' = \psi(V^* \times U')$, and therefore we get the following result.

LEMMA 2. *If $\alpha_0 \in P_-$ then $\mathfrak{l}_0$ is compact and $N \cap \mathfrak{g}_0'$ is connected. On the other hand, if $\alpha_0 \in Q_+ \cup P_0$, $\mathfrak{l}_0$ is not compact and $N \cap \mathfrak{g}_0'$ has exactly three connected components.*

From now on we shall assume that $\alpha_0 \in Q_+ \cup P_0$. Call an element $Z \in \mathfrak{l}_0'$ hyperbolic or elliptic according as $\Omega(Z)$ is positive or negative. Let $\mathfrak{l}_i$ $(i = 1, 2, 3)$ denote the three connected components of $\mathfrak{l}_0'$. We assume that Ω is positive on $\mathfrak{l}_1$ and negative on $\mathfrak{l}_2$ and $\mathfrak{l}_3$ (see Appendix, § 6). Put

$H' = H_{\alpha_0}'$, $X' = X_{\alpha_0}$, $Y' = X_{-\alpha_0}$ if $\alpha_0 \in Q_+$ and $H' = (-1)^{\frac{1}{2}}(X_{\alpha_0} - X_{-\alpha_0})$,

$$X' = \tfrac{1}{2}\{X_{\alpha_0} + X_{-\alpha_0} + (-1)^{\frac{1}{2}} H_{\alpha_0}'\}, \quad Y' = \tfrac{1}{2}\{X_{\alpha_0} + X_{-\alpha_0} - (-1)^{\frac{1}{2}} H_{\alpha_0}'\}$$

if $\alpha_0 \in P_0$. Then $\mathfrak{l}_0 = RH' + RX' + RY'$ and $[H', X'] = 2X'$ $[H', Y']$

$=-2Y'$, $[X', Y'] = H'$. Moreover H' is hyperbolic while $X'-Y'$ is elliptic. We shall suppose that $X'-Y' \in \mathfrak{l}_2$. Let $d\mu_\sigma$ and $d\mu_{\mathfrak{l}}$ denote the Euclidean measures on σ_0 and $\mathfrak{l}_0$, respectively. Every element in $\mathfrak{l}_0$ can be written in the form $tH' + xX' + yY'$ $(t, x, y \in R)$. We normalize $d\mu_{\mathfrak{l}}$ in such a way that[2] $d\mu_{\mathfrak{l}} = (2\pi)^{-1} dt dx dy$. Then since $\mathfrak{z}_0$ is the orthogonal sum of σ_0 and $\mathfrak{l}_0$, we can assume that $dZ = d\mu_\sigma d\mu_{\mathfrak{l}}$. Put

$$\mathfrak{a}_0 = \sigma_0 + RH' \quad \text{and} \quad \mathfrak{b}_0 = \sigma_0 + R(X'-Y').$$

Then $\mathfrak{a}_0$ and $\mathfrak{b}_0$ are Cartan subalgebras of $\mathfrak{z}_0$ (and also of $\mathfrak{g}_0$) which are invariant under θ. Let L be the analytic subgroup of G corresponding to $\mathfrak{l}_0$, and let K_1 be the connected component of 1 in $K \cap L$. Since $\mathfrak{l}_0 \cap \mathfrak{k}_0 = R(X'-Y')$, K_1 is the one-parameter subgroup corresponding to $X'-Y'$. Moreover K_1 is compact. For any $f \in C_c^\infty(\mathfrak{z}_0)$, put

$$\Phi_f(t, H_\sigma) = \int_{-\infty}^{\infty} \bar{f}(H_\sigma + tH' + xX')\,dx \qquad (t > 0),$$

$$\Psi_f(\phi, H_\sigma) = |\phi| \int_0^\infty \bar{f}(H_\sigma + \phi(e^t X' - e^{-t} Y'))\,(e^t - e^{-t})\,dt \quad (\phi \in R, \phi \neq 0)$$

for $H_\sigma \in \sigma_0$. Here $\bar{f}(Z) = \int_{K_1} f(k_1 z)\,dk_1$ $(Z \in \mathfrak{z}_0)$ and dk_1 is the normalized Haar measure on K_1 so that $\int_{K_1} dk_1 = 1$. Let R^+ and R^- denote the sets of all positive and negative real numbers, respectively, and put $R' = R^+ \cup R^-$. Also let $\mathfrak{a}_1$, $\mathfrak{b}_2$, $\mathfrak{b}_3$, respectively, denote the sets of all elements of the form $H_\sigma + tH'$, $H_\sigma + t(X'-Y')$ and $H_\sigma - t(X'-Y')$ $(t \in R^+)$. Put $\mathfrak{b}_0' = \mathfrak{b}_2 \cup \mathfrak{b}_3$. We can also regard Φ_f and Ψ_f as functions on $\mathfrak{a}_1$ and $\mathfrak{b}_0'$, respectively, as follows:

$$\Phi_f(H_\sigma + tH') = \Phi_f(t, H_\sigma), \qquad \Psi_f(H_\sigma + \phi(X'-Y')) = \Psi_f(\phi, H_\sigma)$$

$(H_\sigma \in \sigma_0, t \in R^+, \phi \in R')$. Let $\mathfrak{a}$, $\mathfrak{b}$ be the complexifications of $\mathfrak{a}_0$, $\mathfrak{b}_0$ in $\mathfrak{g}$. Consider the complex analytic subgroup L_c corresponding to $\mathfrak{l}$ in the (connected) complex adjoint group G_c of $\mathfrak{g}$. Define two elements $\nu_{\mathfrak{a}}$ and $\nu_{\mathfrak{b}}$ in L_c as follows. $\nu_{\mathfrak{a}} = 1$, $\nu_{\mathfrak{b}} = \exp\{(-1)^{\frac{1}{2}}(\pi/4)\operatorname{ad}(X_{\alpha_0} + X_{-\alpha_0})\}$ if[2] $\alpha_0 \in Q_+$ and $\nu_{\mathfrak{a}} = \exp\{-(-1)^{\frac{1}{2}}(\pi/4)\operatorname{ad}(X_{\alpha_0} + X_{-\alpha_0})\}$, $\nu_{\mathfrak{b}} = 1$ if $\alpha_0 \in P_0$. Then (see § 6) $H' = \nu_{\mathfrak{a}} H_{\alpha_0}'$, $(X'-Y') = (-1)^{\frac{1}{2}} \nu_{\mathfrak{b}} H_{\alpha_0}'$, and therefore $\nu_{\mathfrak{a}}\mathfrak{h} = \mathfrak{a}$, $\nu_{\mathfrak{b}}\mathfrak{h} = \mathfrak{b}$. For any root α of $\mathfrak{g}$ with respect to $\mathfrak{h}$, we define a root $\alpha_{\mathfrak{a}}$ with respect to $\mathfrak{a}$ by $\alpha_{\mathfrak{a}}(H) = \alpha(\nu_{\mathfrak{a}}^{-1} H)$ $(H \in \mathfrak{a})$. Similarly, define a root $\alpha_{\mathfrak{b}}$ with respect to $\mathfrak{b}$ by $\alpha_{\mathfrak{b}}(H) = \alpha(\nu_{\mathfrak{b}}^{-1} H)$ $(H \in \mathfrak{b})$. We shall say that $\alpha_{\mathfrak{a}}$ and $\alpha_{\mathfrak{b}}$ correspond to α

[2] Here π denotes, as usual, the smallest positive root of the equation $\sin t = 0$.

under $\nu_{\mathfrak{a}}$ and $\nu_{\mathfrak{b}}$, respectively. Put $\tau = (\alpha_0)_{\mathfrak{a}}$ and $\lambda = (\alpha_0)_{\mathfrak{b}}$. Then $\tau(H') = 2$ and $\lambda(X' - Y') = 2(-1)^{\frac{1}{2}}$. Let $P_{\mathfrak{a}}$ and $P_{\mathfrak{b}}$ be the sets of roots with respect to $\mathfrak{a}$ and $\mathfrak{b}$ corresponding to P under $\nu_{\mathfrak{a}}$ and $\nu_{\mathfrak{b}}$, respectively. Define the Euclidean measures $d\mu_{\mathfrak{a}}$, $d\mu_{\mathfrak{b}}$ on $\mathfrak{a}$, $\mathfrak{b}$ by $d\mu_{\mathfrak{a}} = d\mu_\sigma dt$, $d\mu_{\mathfrak{b}} = d\mu_\sigma d\phi$ corresponding to $H = H_\sigma + tH'$ and $\bar{H} = H_\sigma + \phi(X' - Y')$ $(H \in \mathfrak{a}_0;\ \bar{H} \in \mathfrak{b}_0;$ $H_\sigma \in \sigma_0;\ t, \phi \in R)$. Put $\mathfrak{z}_i = \sigma_0 + \mathfrak{l}_i$ $(i = 1, 2, 3)$.

LEMMA 3. *For any $f \in C_c^\infty(\mathfrak{z}_0)$ define Φ_f, Ψ_f as above. Then Φ_f and Ψ_f lie in $\mathcal{C}(\mathfrak{a}_1)$ and $\mathcal{C}(\mathfrak{b}_0')$, respectively, and*

$$\int_{\mathfrak{z}_1} f(Z)\,dZ = \tfrac{1}{2}\int_{\mathfrak{a}_1} \tau\Phi_f d\mu_{\mathfrak{a}}, \qquad \int_{\mathfrak{z}_i} f(Z)\,dZ = \tfrac{1}{2}\int_{\mathfrak{b}_i} |\lambda|\,\Psi_f d\mu_{\mathfrak{b}} \quad (i = 2, 3).$$

Since σ is the center of $\mathfrak{z}$, it follows without difficulty that $\partial(p)\Phi_f = \Phi_{\partial(p)f}$, $\partial(p)\Psi_f = \Psi_{\partial(p)f}$ for $p \in S(\sigma)$ (see [4(d), § 2] for notation). Therefore since the carrier of f is compact, our statement follows from the work of the Appendix, especially Lemma 16.

Let Ξ' be the normalizer of σ_0 in G and Ξ_0 the analytic subgroup of G corresponding to $\mathfrak{z}_0$.

LEMMA 4. *Ξ' is also the normalizer of $\mathfrak{z}_0$ in G and every element of Ξ' leaves $\mathfrak{l}_0$ invariant. Moreover, $\Xi' \supset \Xi \supset \Xi_0$ and Ξ_0 is of finite index in Ξ'.*

Since σ_0 is the center of $\mathfrak{z}_0$ and $\mathfrak{z}_0$ the centralizer of σ_0 in $\mathfrak{g}_0$, it is obvious that Ξ' is also the normalizer of $\mathfrak{z}_0$ in G. Moreover, since $\mathfrak{l}_0$ is the derived algebra of $\mathfrak{z}_0$, $x\mathfrak{l}_0 = \mathfrak{l}_0$ for $x \in \Xi'$. Finally, since $\mathfrak{z}_0$ is the centralizer of H_0 in $\mathfrak{g}_0$, it is clear that $\Xi_0 \subset \Xi \subset \Xi'$. The finiteness of Ξ'/Ξ_0 is a consequence of Lemma 15 of the Appendix (Section 6).

We shall now show that if U is sufficiently small, $xU \cap U = \emptyset$ $(x \in G)$ unless $x \in \Xi$. This is done as follows. Since $S(\mathfrak{l}) \subset S(\mathfrak{g})$, we can regard Ω as a polynomial function on $\mathfrak{g}$. But then, since σ and $\mathfrak{l}$ are orthogonal, $\Omega(H + Z) = \Omega(Z)$ $(H \in \sigma, Z \in \mathfrak{l})$, and therefore $\Omega(H_0) = 0$. Choose a positive number δ. Then we may assume that U is so small that $|\Omega(Z)| < \delta^2 |\alpha(H_\sigma(Z))|^2$ $(Z \in U)$ for all $\alpha \neq \alpha_0$ in P. Here $H_\sigma(Z)$ denotes the orthogonal projection of Z in σ for any $Z \in \mathfrak{z}$. Since $[\sigma, \mathfrak{l}] = \{0\}$, it is clear that $H_\sigma(yZ) = H_\sigma(Z)$ $(y \in L_c, Z \in \mathfrak{z})$. Hence

$$\min_{\substack{\alpha \neq \alpha_0 \\ \alpha \in P}} |\alpha(H_\sigma(Z))| = \min_{\substack{\alpha \neq \tau \\ \alpha \in P_{\mathfrak{a}}}} |\alpha(H_\sigma(Z))| = \min_{\substack{\alpha \neq \lambda \\ \alpha \in P_{\mathfrak{b}}}} |\alpha(H_\sigma(Z))|.$$

Moreover, if $Z \in \mathfrak{a}$, $\Omega(Z) = 2\tau(Z)^2$ (see Section 6). Let $\tilde{U} = \bigcup_{y \in L} yU$. Then it follows that if $Z \in \tilde{U} \cap \mathfrak{a}_0$

$$|\tau(Z)| < \delta\,|\alpha(H_\sigma(Z))| \text{ for all } \alpha \neq \tau \text{ in } P_{\mathfrak{a}}.$$

Similarly, if $Z \in \tilde{U} \cap \mathfrak{b}_0$, then $|\lambda(Z)| < \delta |\alpha(H_\sigma(Z))|$ for all $\alpha \neq \lambda$ in $P_{\mathfrak{b}}$. Now suppose δ is so small that $\delta \leqq \frac{1}{2}$ and

$$\delta |\alpha(H')| \leqq 1 \qquad (\alpha \in P_{\mathfrak{a}}), \qquad \delta |\beta(X'-Y')| \leqq 1 \qquad (\beta \in P_{\mathfrak{b}}).$$

Then since $Z = H_\sigma(Z) + \frac{1}{2}\tau(Z)H'$ for $Z \in \tilde{U} \cap \mathfrak{a}_0$, it follows that $\alpha(Z) = \alpha(H_\sigma(Z)) + \frac{1}{2}\tau(Z)\alpha(H')$ for $\alpha \in P_{\mathfrak{a}}$. Therefore if $\alpha \neq \tau$,

$$|\alpha(Z)| \geqq 2\delta |\alpha(Z)| \geqq 2\delta\{|\alpha(H_\sigma(Z))| - \tfrac{1}{2}|\tau(Z)||\alpha(H')|\}$$
$$> (2-\delta|\alpha(H')|)|\tau(Z)| \geqq |\tau(Z)|.$$

This proves that $|\alpha(Z)| > |\tau(Z)|$ for $Z \in \tilde{U} \cap \mathfrak{a}_0$ and $\alpha \neq \tau$ in $P_{\mathfrak{a}}$. Similarly, $|\alpha(Z)| > |\lambda(Z)|$ $(\alpha \in P_{\mathfrak{b}}, \alpha \neq \lambda)$ for $Z \in \tilde{U} \cap \mathfrak{b}_0$. Moreover, we have the following result.

LEMMA 5. *The two Cartan subalgebras $\mathfrak{a}_0$ and $\mathfrak{b}_0$ are not conjugate under G.*

Since $\theta(\sigma_0) = \sigma_0$, $H' \in \mathfrak{p}_0$ and $X' - Y' \in \mathfrak{k}_0$ (see [4(e), § 12]), it is clear that $\mathfrak{a}_0 \cap \mathfrak{k}_0 = \sigma_0 \cap \mathfrak{k}_0$ while $\mathfrak{b}_0 \cap \mathfrak{k}_0 = \sigma_0 \cap \mathfrak{k}_0 + R(X'-Y')$. Hence $l_-(\mathfrak{b}_0) = l_-(\mathfrak{a}_0) + 1$ in the notation of [4(e), § 8]. Therefore $\mathfrak{a}_0$ and $\mathfrak{b}_0$ cannot be conjugate.

Put $U_i = U \cap \mathfrak{z}_i$ $(i = 1,2,3)$.

COROLLARY. *Suppose $Z_1 \in U_1$ and $Z_2 \in U_2 \cup U_3$. Then Z_1 and Z_2 cannot be conjugate under G.*

For otherwise suppose $Z_2 = xZ_1$ $(x \in G)$. Since $U_1 \subset \sigma_0 + \mathfrak{l}_1$, we can select (see Section 6) $y_1 \in L$ such that $Z_1' = y_1 Z_1 \in \mathfrak{a}_1$. Similarly, we can choose $y_2 \in L$ such that $Z_2' = y_2 Z_2 \in \mathfrak{b}_0'$. Obviously Z_1, Z_2 are regular in $\mathfrak{g}_0$, and therefore the same holds for Z_1', Z_2'. Hence $\mathfrak{a}_0$ and $\mathfrak{b}_0$, respectively, are the centralizers of Z_1' and Z_2' in $\mathfrak{g}_0$. Moreover, $Z_2' = x'Z_1'$ $(x' = y_2 x y_1^{-1})$, and therefore $\mathfrak{b}_0 = x'\mathfrak{a}_0$. As this contradicts Lemma 5, the corollary follows.

Now suppose $xZ \in U$ for some $x \in G$ and $Z \in U$. First, let us assume that $Z \in U_1$. Then obviously xZ is regular in $\mathfrak{g}_0$, and so it follows from the Corollary to Lemma 5 that $xZ \in U_1$. Hence $y_1 Z$, $y_2 xZ$ are in $U_\sigma + R^+H'$ for some $y_1, y_2 \in L$. Put $Z_1 = y_1 Z$ and $x_1 = y_2 x y_1^{-1}$. Then since $\mathfrak{a}_0$ is the centralizer in $\mathfrak{g}_0$ both of Z_1 and $x_1 Z_1$, we conclude that $x_1\mathfrak{a}_0 = \mathfrak{a}_0$. Let $W_{\mathfrak{a}}$ be the Weyl group of $\mathfrak{g}$ with respect to $\mathfrak{a}$. Then x_1 defines an element $s \in W_{\mathfrak{a}}$. Put $s\tau = \tau'$ and $Z_2 = x_1 Z_1$. It is clear that $\tau'(Z_2) = \tau(Z_1)$. On the other hand, since $Z_1 \in \tilde{U} \cap \mathfrak{a}_0$, $|\tau(Z_1)| = \min_{\alpha \in P_{\mathfrak{a}}} |\alpha(Z_1)|$. Therefore $|\tau'(Z_2)| = \min_{\alpha \in P_{\mathfrak{a}}} |\alpha(Z_2)|$. But since Z_2 is also in $\tilde{U} \cap \mathfrak{a}_0$, we know that $|\tau(Z_2)|$

$< |\alpha(Z_2)|$ for any root $\alpha \neq \tau$ in $P_{\mathfrak{a}}$. Hence $\tau' = \pm \tau$. Now σ_0 is the set of all $H \in \mathfrak{a}_0$ where $\tau(H) = 0$. Therefore $x_1\sigma_0 = s\sigma_0 = \sigma_0$. Since L is contained in Ξ, this proves that $x \in \Xi'$.

Similarly, if $Z \in U_2 \cup U_3$, we can choose $y_1 \in L$ such that $Z_1 = y_1 Z \in U_\sigma + R'(X' - Y')$. Again it follows from the Corollary to Lemma 5 that $xZ \in U_2 \cup U_3$, and therefore $y_2 x Z \in U_\sigma + R'(X' - Y')$ for some $y_2 \in L$. Hence $x_1\mathfrak{b}_0 = \mathfrak{b}_0$ $(x_1 = y_2 x y_1^{-1})$ and the corresponding element s of the Weyl group $W_{\mathfrak{b}}$ (of $\mathfrak{g}$ with respect to $\mathfrak{b}$) maps λ into $\pm\lambda$, and therefore again $x \in \Xi'$.

Finally suppose Z is singular. For any $X \in \mathfrak{g}_0$, let $\mathfrak{n}_X$ denote the set of all $Y \in \mathfrak{g}_0$ such that $(\mathrm{ad}\, X)^q Y = 0$ for some integer $q \geqq 1$. We call $\mathfrak{n}_X$ the null space of X. It is obvious that if X is a singular element in U, $\mathfrak{n}_X = \mathfrak{z}_0$. Hence $\mathfrak{z}_0$ is the null space of both Z and xZ, and therefore $x\mathfrak{z}_0 = \mathfrak{z}_0$. In view of Lemma 4 this implies again that $x \in \Xi'$.

Now from Lemma 4, we can select a finite number of elements $u_0 = 1, u_1, \cdots, u_M \in \Xi'$ such that the cosets $u_i\Xi_0$ $(0 \leqq i \leqq M)$ are distinct and their union is Ξ'. Suppose $u_i \in \Xi$ for $0 \leqq i \leqq m$ and $u_i \notin \Xi$ for $m < i \leqq M$. Then $u_i H_0 \neq H_0$ $(m < i \leqq M)$, and we may assume that U_σ is so small that $U_\sigma \cap u_i U_\sigma = \emptyset$ $(m < i \leqq M)$. Select i such that $x \in u_i \Xi_0$. Then we claim that $i \leqq m$. For otherwise, suppose $i > m$. Since $U = U_\sigma + U_{\mathfrak{l}}$, it follows that $xU = xU_\sigma + xU_{\mathfrak{l}}$. On the other hand, from Lemma 4, xU_σ and $xU_{\mathfrak{l}}$ are contained in σ_0 and $\mathfrak{l}_0$, respectively, and since σ is the center of $\mathfrak{z}$, $xU_\sigma = u_i U_\sigma$, and therefore $U_\sigma \cap xU_\sigma = \emptyset$. But obviously, this implies that $U \cap xU = \emptyset$, and we get a contradiction with our assumption that $xZ \in U \cap xU$. Hence $i \leqq m$ and therefore $x \in \Xi$.

Put $\bar{U}_\sigma = \bigcap_{1 \leqq i \leqq m} u_i U_\sigma$, $\bar{U}_{\mathfrak{l}} = \bigcup_{\xi \in \Xi} \xi U_{\mathfrak{l}}$ and $\bar{U} = \bar{U}_\sigma + \bar{U}_{\mathfrak{l}}$. Since U_σ is convex, the same holds for $\bar{U}_\sigma$. Moreover, in view of its definition, Ω is invariant under any automorphism of $\mathfrak{l}$. Hence if $Z \in \bar{U}_{\mathfrak{l}}$, it is clear that $|\Omega(Z)| < \delta^2 |\alpha(H_\sigma(Z))|^2$ for all roots $\alpha \neq \alpha_0$ in P. Therefore it follows from the above proof that $x\bar{U} \cap \bar{U} = \emptyset$ $(x \in G)$ unless $x \in \Xi$.

LEMMA 6. *It is possible to select U in Lemma 1 so that it satisfies the following two further conditions.*

(1) $\xi U = U$ $(\xi \in \Xi)$,

(2) $U \cap xU = \emptyset$ *for any $x \in G$ which is not in Ξ.*

As we have already seen, $\bar{U}$ satisfies the above two conditions. The univalence of the mapping $(x^*, Z) \to \phi(x^*)Z$ of $V^* \times \bar{U}$ into $\mathfrak{g}_0$ is an immediate consequence of the second condition. Since the earlier conditions

imposed on U (see p. 656) are obviously fulfilled by $\bar{U}$, the statement of the lemma follows.

Let A and B be the Cartan subgroups of G corresponding to $\mathfrak{a}_0$ and $\mathfrak{b}_0$, respectively. Since L is a normal subgroup of Ξ' (see Lemma 4), AL and BL are subgroups of Ξ and $AL = LA$, $BL = LB$. Let $[\Xi : LB]$ denote the index of LB in Ξ.

LEMMA 7. *$\Xi = LA$ and $[\Xi : LB] \leqq 2$. Let s_λ denote the Weyl reflexion in $W_{\mathfrak{b}}$ corresponding to the root $\lambda \in P_{\mathfrak{b}}$. Then $[\Xi : LB] = 2$ if and only if there exists an element $\xi \in \Xi$ which coincides with s_λ on $\mathfrak{b}$.*

Let x be any element in Ξ. Then $x\mathfrak{l}_0 = \mathfrak{l}_0$ from Lemma 4, and therefore $\Omega(xH') = \Omega(H')$, $\Omega(x(X' - Y')) = \Omega(X' - Y')$. This shows that $xH' \in \mathfrak{l}_1$ and $x(X' - Y') \in \mathfrak{l}_2 \cup \mathfrak{l}_3$. Hence we can select $y_1, y_2 \in L$ such that $y_1^{-1}xH' \in R^+H'$ and $y_2^{-1}x(X' - Y') \in R'(X' - Y')$. But since $\mathrm{sp}(\mathrm{ad}(zX))^2 = \mathrm{sp}(\mathrm{ad}\,X)^2$ $(z \in G, X \in \mathfrak{g})$ this implies that $y_2^{-1}xH' = H'$ and $y_2^{-1}x(X' - Y') = \pm(X' - Y')$. Now if t is sufficiently small and positive, $H_0 + tH'$ is nonsingular and so its centralizer in G is A (see [4(c), Lemma 3]). Hence $y_1^{-1}x \in A$, and so $x \in LA$. This proves that $\Xi = LA$. On the other hand, choose, if possible, an element $\xi \in \Xi$ such that $\xi(X' - Y') = -(X' - Y')$. In case $y_2^{-1}x(X' - Y') = X' - Y'$, the above argument applied to $H_0 + t(X' - Y')$ $(t \in R')$ shows that $x \in LB$. So now suppose $y_2^{-1}x(X' - Y') = -(X' - Y')$. Then $\xi^{-1}y_2^{-1}x(X' - Y') = (X' - Y')$, and therefore $\xi^{-1}y_2^{-1}x \in B$. This proves that $x \in L\xi B = \xi LB$ in this case. Hence $[\Xi : LB] \leqq 2$. Moreover, it is obvious that $[\Xi : LB] = 1$ if ξ does not exist. So let us assume that ξ exists. We claim $\xi \notin LB$. For otherwise, $\xi = yb$ $(y \in L, b \in B)$, and therefore $y(X' - Y') = \xi(X' - Y') = -(X' - Y')$. But this is impossible (see Section 6), and therefore $[\Xi : LB] = 2$ in this case. Since $\xi\sigma = \sigma$, it is clear that $\xi\mathfrak{b} = \mathfrak{b}$. Let s be the element of $W_{\mathfrak{b}}$ corresponding to ξ. Then

$$\begin{aligned} s(H_0 + t(X' - Y')) &= \xi(H_0 + t(X' - Y')) \\ &= H_0 - t(X' - Y') = s_\lambda(H_0 + t(X' - Y')) \quad (t \in R). \end{aligned}$$

But if t is suitably chosen, $H_0 + t(X' - Y')$ is regular in $\mathfrak{g}_0$. Then since $s^{-1}s_\lambda$ leaves it fixed, we can conclude (see [4(c), Lemma 4]) that $s = s_\lambda$. This completes the proof of the lemma.

Now select U once for all according to Lemma 6.

COROLLARY. *Let x be an element in G such that $x(\mathfrak{a}_1 \cap U)$ meets $\mathfrak{a}_1 \cap U$. Then $x \in A$. Similarly, if $y(\mathfrak{b}_i \cap U)$ meets $\mathfrak{b}_i \cap U$ $(y \in G, i = 2, 3)$, then $y \in B$.*

Select $H_1 \in \mathfrak{a}_1 \cap U$ such that $xH_1 \in \mathfrak{a}_1 \cap U$. Then since $\mathfrak{a}_0$ is the centralizer in $\mathfrak{g}_0$ of both H_1 and $H_2 = xH_1$, $\mathfrak{a}_0 = x\mathfrak{a}_0$. Hence x defines an element $s \in W_\mathfrak{a}$. But $x\sigma = \sigma$, since $x \in \Xi$ from Lemma 6. Therefore it is clear that $s\tau = \pm \tau$. But $\tau(s^{-1}H_2) = \tau(H_1) > 0$ and $H_2 \in \mathfrak{a}_1$. So we conclude that $s\tau = \tau$, and therefore $sH' = H'$. This proves that $x(H_0 + tH') = H_0 + tH'$ $(t \in R)$, and therefore from [4(c), Lemma 3] $x \in A$. The proof in the other case is quite similar.

Define $N = \psi(V^* \times U)$ as in Lemma 1, and put $U_i = U \cap \mathfrak{z}_i$, $N_i = \psi(V^* \times U_i)$ $(i = 1,2,3)$. Moreover, let $\mathfrak{b}^+ = \mathfrak{b}_2$, $\mathfrak{b}^- = \mathfrak{b}_3$, $N^+ = N_2$, $N^- = N_3$ if $[\Xi : LB] = 1$ and $\mathfrak{b}^+ = \mathfrak{b}^- = \mathfrak{b}_2 \cup \mathfrak{b}_3$, $N^+ = N^- = N_2 \cup N_3$ if $[\Xi : LB] = 2$. Let p be any polynomial in $S(\mathfrak{g})$ which is invariant under L. Then we denote by $p_\mathfrak{a}$ and $p_\mathfrak{b}$ its restrictions on $\mathfrak{a}$ and $\mathfrak{b}$, respectively. The following lemma contains the main result of this section.

Lemma 8. *For any $\gamma \in C_c^\infty(U)$, define $f_\gamma \in C_c^\infty(N)$ as in Lemma* 1. *Then*

$$\int_{N_1} p\partial(q) f_\gamma \, dX = \tfrac{1}{2} \int_{\mathfrak{a}_1} \tau p_\mathfrak{a} \mid \zeta_\mathfrak{a} \mid^{\frac{1}{2}} \partial(q_\mathfrak{a}) \Phi_{|\zeta|^{\frac{1}{2}}\gamma} d\mu_\mathfrak{a},$$

$$\int_{N^\pm} p\partial(q) f_\gamma \, dX = \tfrac{1}{2} \int_{\mathfrak{b}^\pm} \mid \lambda \mid p_\mathfrak{b} \mid \zeta_\mathfrak{b} \mid^{\frac{1}{2}} \partial(q_\mathfrak{b}) \Psi_{|\zeta|^{\frac{1}{2}}\gamma} \, d\mu_\mathfrak{b}$$

for $p, q \in I(\mathfrak{g})$ and $\gamma \in C_c^\infty(U)$.

If U_σ is the orthogonal projection of U in σ_0, $U \cap \mathfrak{a}_1 \subset U_\sigma + R_1 H' \subset \mathfrak{a}_0 \cap \mathfrak{g}_0'$ where R_1 is an open interval in R^+. Therefore, since U_σ is connected, $U \cap \mathfrak{a}_1$ is contained in a connected component of $\mathfrak{a}_0 \cap \mathfrak{g}_0'$. Put $\pi_\mathfrak{a} = \prod_{\alpha \in P_\mathfrak{a}} \alpha$. Then $\pi_\mathfrak{a}^2$ takes only real values on $\mathfrak{a}_0$, and so we can select a complex number ϵ such that $\epsilon^4 = 1$ and $\epsilon\pi_\mathfrak{a}$ is everywhere positive on $U \cap \mathfrak{a}_1$. Since $U_1 \subset U_\sigma + \mathfrak{l}_1$, it is obvious that $N_1 \subset \bigcup_{x \in G} x(\mathfrak{a}_1 \cap U)$. We claim $N_1 = N \cap \bigcup_{x \in G} x(\mathfrak{a}_1 \cap U)$. For suppose $zH \in N$ for some $z \in G$ and $H \in \mathfrak{a}_1 \cap U$. Then $zH = \phi(x^*)Z$ for some $x^* \in V^*$ and $Z \in U$. It follows from the Corollary to Lemma 5 that $Z \in U_1$ and therefore $Z = yH_1$ $(y \in L, H_1 \in \mathfrak{a}_1 \cap U)$. But then we conclude from Lemma 6 that $z^{-1}\phi(x^*)y \in \Xi$, and therefore $z \in \phi(x^*)\Xi$. Since $\Xi H \subset U_1$ (by the Corollary to Lemma 5), it follows that $zH \in \phi(x^*)U_1 \subset N_1$, and this proves our assertion. *Moreover, the above argument shows that if $zH \in N$ $(z \in G, H \in \mathfrak{a}_1 \cap U)$, then $z^* \in V^*$.* We shall need this fact presently.

Put $\bar{G} = G/A$ and define $\bar{x}H = xH$ $(x \in G, H \in \mathfrak{a}_0)$, where $x \rightarrow \bar{x}$ is the natural mapping of G onto $\bar{G}$. Then in view of the above result and the

Corollary to Lemma 7, we can normalize the invariant measure $d\bar{x}$ on $\bar{G}$ in such a way (see [4(a), p. 501]) that

$$\int_{N_1} g(X)\,dX = \int_{\mathfrak{a}_1\cap U} (\epsilon\pi_{\mathfrak{a}}(H))^2 d\mu_{\mathfrak{a}} \int_{\bar{G}} g(\bar{x}H)\,d\bar{x} \qquad (g \in C_c(N_1)).$$

Put $\bar{\Xi} = \Xi/A$ and let $d\bar{\xi}$ denote the invariant measure on $\bar{\Xi}$. Then it is clear that if $d\bar{\xi}$ is suitably normalized,

$$\int_{\bar{G}} g(\bar{x}H)\,d\bar{x} = \int_{G^*} dx^* \int_{\bar{\Xi}} g(x(\bar{\xi}H))\,d\bar{\xi} \qquad (g \in C_c^\infty(\mathfrak{g}_0), H \in \mathfrak{a}_0 \cap \mathfrak{g}_0').$$

Now put $\psi_g(H) = \pi_{\mathfrak{a}}(H) \int_{\bar{G}} g(\bar{x}H)\,d\bar{x}$ $(H \in \mathfrak{a}_0 \cap \mathfrak{g}_0')$ for $g \in C_c^\infty(\mathfrak{g}_0)$. (It follows from Theorem 3 of [4(e)] that the above integral is convergent.) Then if $g = f_\gamma$ $(\gamma \in C_c^\infty(U))$,

$$\psi_{f_\gamma}(H) = \pi_{\mathfrak{a}}(H) \int_{G^*} dx^* \int_{\bar{G}} f_\gamma(x(\bar{\xi}H))\,d\bar{\xi} \qquad (H \in \mathfrak{a}_1 \cap U).$$

We have seen above that if $zH \in N$ for some $z \in G$ and $H \in \mathfrak{a}_1 \cap U$, then $z^* \in V^*$. Therefore

$$\psi_{f_\gamma}(H) = \pi_{\mathfrak{a}}(H) \int_{V^*} dx^* \int_{\bar{G}} f_\gamma(\phi(x^*)(\bar{\xi}H))\,d\bar{\xi} \qquad (H \in \mathfrak{a}_1 \cap U).$$

But $f_\gamma(\phi(x^*)Z) = a_0(x^*)\gamma(Z)$ $(x^* \in V^*, Z \in U)$ from the definition of f_γ. Therefore, since U is invariant under Ξ, we get

$$\begin{aligned}\psi_{f_\gamma}(H) &= \pi_{\mathfrak{a}}(H) \int_{V^*} a_0(x^*)\,dx^* \int_{\bar{\Xi}} \gamma(\bar{\xi}H)\,d\bar{\xi} \\ &= \pi_{\mathfrak{a}}(H) \int_{\bar{\Xi}} \gamma(\bar{\xi}H)\,d\bar{\xi} \qquad (H \in \mathfrak{a}_1 \cap U).\end{aligned}$$

Moreover, since $\Xi = LA$ it is obvious (see Section 6) that

$$\tau(H) \int_{\bar{\Xi}} \gamma(\bar{\xi}H)\,d\bar{\xi} = c\Phi_\gamma(H) \qquad (H \in \mathfrak{a}_1),$$

where c is a positive number independent of H or γ. Therefore

$$\psi_{f_\gamma}(H) = c\pi_{\mathfrak{a}}'(H)\Phi_\gamma(H) \qquad (\gamma \in C_c^\infty(U), H \in \mathfrak{a}_1 \cap U),$$

where $\pi_{\mathfrak{a}}'$ is the product of all roots $\alpha \neq \tau$ in $P_{\mathfrak{a}}$. On the other hand, ζ is invariant under L and since τ is everywhere positive on $\mathfrak{a}_1$, it is clear that $|\zeta_{\mathfrak{a}}|^{\frac{1}{2}} = \epsilon\pi_{\mathfrak{a}}'$ on $\mathfrak{a}_1 \cap U$. Hence $\psi_{f_\gamma} = c\epsilon^{-1}\Phi_{|\zeta|^{\frac{1}{2}}\gamma}$ on $\mathfrak{a}_1 \cap U$. Now suppose $p, q \in I(\mathfrak{g})$. Then

$$\int_{N_1} p(X) f_\gamma(X;\partial(q))\,dX = \int_{\mathfrak{a}_1\cap U} (\epsilon\pi_{\mathfrak{a}}(H))^2 p(H)\,d\mu_{\mathfrak{a}} \int_{\bar{G}} f_\gamma(\bar{x}H;\partial(q))\,d\bar{x}$$
$$= \epsilon^2 \int_{\mathfrak{a}_1\cap U} \pi_{\mathfrak{a}}(H) p(H) \psi_{f\gamma}(H;\partial(q_{\mathfrak{a}}))\,d\mu_{\mathfrak{a}}$$

from Theorem 3 of [4(e)]. But

$$\epsilon^2\pi_{\mathfrak{a}}\partial(q_{\mathfrak{a}})\psi_{f\gamma} = c\epsilon\pi_{\mathfrak{a}}'\tau\partial(q_{\mathfrak{a}})(\Phi_{|\zeta|^{\frac{1}{2}}\gamma}) = c\,|\,\zeta_{\mathfrak{a}}\,|^{\frac{1}{2}}\tau\partial(q_{\mathfrak{a}})(\Phi_{|\zeta|^{\frac{1}{2}}\gamma})$$

on $\mathfrak{a}_1 \cap U$. Therefore since $\Phi_{|\zeta|^{\frac{1}{2}}\gamma}$ is zero on $\mathfrak{a}_1$ outside $\mathfrak{a}_1 \cap U$, it follows that

$$\int_{N_1} p(X) f_\gamma(X;\partial(q))\,dX = c\int_{\mathfrak{a}_1} |\,\zeta_{\mathfrak{a}}\,|^{\frac{1}{2}}\tau p_{\mathfrak{a}}\partial(q_{\mathfrak{a}})\Phi_{|\zeta|^{\frac{1}{2}}\gamma}\,d\mu_{\mathfrak{a}}.$$

In order to determine c, we put $p = q = 1$. Then

$$\int_{N_1} f_\gamma\,dX = c\int_{\mathfrak{a}_1} |\,\zeta_{\mathfrak{a}}\,|^{\frac{1}{2}}\tau\Phi_{|\zeta|^{\frac{1}{2}}\gamma}\,d\mu_{\mathfrak{a}} = c\int_{\mathfrak{a}_1} \tau\Phi_{|\zeta|\gamma}\,d\mu_{\mathfrak{a}} = 2c\int_{U_1} |\,\zeta\,|\,\gamma\,dZ$$

from Lemma 3. Therefore we conclude from Lemma 1 that $c = \frac{1}{2}$.

Now we come to the second formula. Since $U_i \subset U_\sigma + \mathfrak{l}_i$ $(i = 2, 3)$, we prove as before that $N_i \subset \bigcup_{x\in G} x(\mathfrak{b}_i \cap U)$. Moreover, by the same argument as used above, we show that $N_2 \cup N_3 = N \cap (\bigcup_{x\in G} x(\mathfrak{b}_0' \cap U))$. Now first suppose $[\Xi : LB] = 1$. Then no point in N_2 can be conjugate (under G) to a point in N_3. For otherwise, we could choose $H_i \in \mathfrak{b}_i \cap U$ $(i = 2, 3)$ and $x \in G$ such that $xH_2 = H_3$. But as we have seen during the proof of Lemma 7, this implies that x coincides with s_λ on $\mathfrak{b}$, which is impossible since $[\Xi : LB] = 1$. Hence $N_i \cap (\bigcup_{x\in G} x(\mathfrak{b}_j \cap U))$ is empty if $i = 2$, $j = 3$ or $i = 3$, $j = 2$. This shows that $N_i = N \cap \bigcup_{x\in G} x(\mathfrak{b}_i \cap U)$ $(i = 2, 3)$ in this case, and the proof of the second formula now proceeds in the same way as that of the first.

So let us now assume that $[\Xi : LB] = 2$ and select an element $z \in \Xi$ which coincides with s_λ on $\mathfrak{b}$. Then $z\mathfrak{b}_2 = \mathfrak{b}_3$ and therefore $N_2 \cup N_3 = N \cap (\bigcup_{x\in G} x(\mathfrak{b}_2 \cap U))$. Define $\pi_{\mathfrak{b}} = \prod_{\alpha\in P_{\mathfrak{b}}} \alpha$ and choose, as before, a complex number ϵ such that $\epsilon^4 = 1$ and $\epsilon\pi_{\mathfrak{b}}^2$ is everywhere positive on $U \cap \mathfrak{b}_2$. We use our earlier notation except that now $\bar{G} = G/B$, $\bar{\Xi} = \Xi/B$ and

$$\psi_g(H) = \pi_{\mathfrak{b}}(H)\int_{\bar{G}} g(\bar{x}H)\,d\bar{x} \qquad (H \in \mathfrak{b}_0 \cap \mathfrak{g}_0', g \in C_c^\infty(\mathfrak{g}_0)).$$

Then again

$$\psi_{f\gamma}(H) = \pi_{\mathfrak{b}}(H)\int_{\bar{\Xi}} \gamma(\bar{\xi}H)\,d\bar{\xi} \qquad (H \in \mathfrak{b}_2 \cap \mathfrak{g}_0', \gamma \in C_c^\infty(U)).$$

Let $\bar{L}=(LB)/B$. Then since $\Xi=LB\cup(LB)z$, it follows that

$$\int_{\bar{\Xi}}\gamma(\bar{\xi}H)\,d\bar{\xi}=\int_{\bar{L}}\gamma(\bar{\xi}H)\,d\bar{\xi}+\int_{\bar{L}}\gamma(\bar{\xi}(s_\lambda H))\,d\bar{\xi}\qquad (H\in\mathfrak{b}_2).$$

Moreover it is clear (see Section 6) that there exists a positive number c such that

$$|\lambda(H)|\int_{\bar{L}}\gamma(\bar{\xi}H)\,d\bar{\xi}=c\Psi_\gamma(H)\qquad (H\in\mathfrak{b}_2\cup\mathfrak{b}_3,\gamma\in C_c^\infty(U)).$$

For any function g on $\mathfrak{b}_2\cup\mathfrak{b}_3$, let g^{s_λ} denote the function $H\to g(s_\lambda H)$ $(H\in\mathfrak{b}_2\cup\mathfrak{b}_3)$. Then since $\epsilon\pi_\mathfrak{b}=|\pi_\mathfrak{b}|=|\zeta_\mathfrak{b}|^{\frac{1}{2}}|\lambda|$ on $\mathfrak{b}_2\cap U$, we conclude that

$$\psi_{f\gamma}=c\epsilon^{-1}\{\Psi_{|\zeta|^{\frac{1}{2}}\gamma}+(\Psi_{|\zeta|^{\frac{1}{2}}\gamma})^{s_\lambda}\}$$

on $\mathfrak{b}_2\cap U$. Now suppose $p,q\in I(\mathfrak{g})$. Then since

$$N^+=N_2\cup N_3=N\cap\bigcup_{x\in G}x(\mathfrak{b}_2\cap U),$$

it follows that

$$\int_{N^+}p(X)f_\gamma(X;\partial(q))\,dX=\int_{\mathfrak{b}_2\cap U}(\epsilon\pi_\mathfrak{b}(H))^2p(H)\,d\mu_\mathfrak{b}\int_{\bar{G}}f_\gamma(\bar{x}H;\partial(q))\,d\bar{x}$$

$=\epsilon^2\int_{\mathfrak{b}_2\cap U}\pi_\mathfrak{b}(H)p(H)\psi_{f\gamma}(H;\partial(q_\mathfrak{b}))\,d\mu_\mathfrak{b}$ as before. Furthermore,

$$\epsilon^2\pi_\mathfrak{b}\partial(q_\mathfrak{b})\psi_{f\gamma}=c\,|\zeta_\mathfrak{b}|^{\frac{1}{2}}|\lambda|\,\partial(q_\mathfrak{b})\{\Psi_{|\zeta|^{\frac{1}{2}}\gamma}+(\Psi_{|\zeta|^{\frac{1}{2}}\gamma})^{s_\lambda}\}$$

on $\mathfrak{b}_2\cap U$. Therefore, since $\mathfrak{b}^+=\mathfrak{b}_2\cup\mathfrak{b}_3$, it is clear that

$$\int_{N^+}p\partial(q)f_\gamma\,dX=c\int_{\mathfrak{b}^+}|\zeta_\mathfrak{b}|^{\frac{1}{2}}|\lambda|\,p_\mathfrak{b}\partial(q_\mathfrak{b})\Psi_{|\zeta|^{\frac{1}{2}}\gamma}\,d\mu_\mathfrak{b}.$$

In order to determine c, we again put $p=q=1$. Then

$$\int_{N^+}f_\gamma\,dX=c\int_{\mathfrak{b}^+}|\lambda|\,\Psi_{|\zeta|\gamma}\,d\mu_\mathfrak{b}=2c\int_{U_2\cup U_3}|\zeta|\,\gamma\,dZ$$

from Lemma 3. Therefore $c=\frac{1}{2}$ from Lemma 1. This completes the proof of Lemma 8.

Define the polynomial $\eta\in I(\mathfrak{g})$ as in [4(e), § 6] and put $\epsilon_\mathfrak{a}=|\pi_\mathfrak{a}|\pi_\mathfrak{a}^{-1}$, $\epsilon_\mathfrak{b}=|\pi_\mathfrak{b}|\pi_\mathfrak{b}^{-1}$ on $\mathfrak{a}_1\cap U$ and $\mathfrak{b}_0'\cap U$, respectively. Then as we have seen above, $\epsilon_\mathfrak{a}$ and $\epsilon_\mathfrak{b}$ are constant on $\mathfrak{a}_1\cap U$ and $\mathfrak{b}_i\cap U$ $(i=2,3)$, respectively. Moreover since $\pi_\mathfrak{b}^{s_\lambda}=-\pi_\mathfrak{b}$, it is clear that $\epsilon_\mathfrak{b}(H_2)=-\epsilon_\mathfrak{b}(H_3)$ if $H_i\in b_i\cap U$ $(i=2,3)$.

COROLLARY. *Let k be any integer $\geqq 0$. Then for any $\xi \in I(\mathfrak{g})$,*

$$\int_{N_1} \eta^k \partial(\eta^k \xi) f_\gamma \, dX = \tfrac{1}{2} \int_{\mathfrak{a}_1} \epsilon_\mathfrak{a} \pi_\mathfrak{a}^{2k+1} \partial(\pi_\mathfrak{a}^{2k} \xi_\mathfrak{a}) \Phi_{|\zeta|^{\frac{1}{2}}\gamma} d\mu_\mathfrak{a},$$

$$\int_{N^\pm} \eta^k \partial(\eta^k \xi) f_\gamma \, dX = \tfrac{1}{2} \int_{\mathfrak{b}^\pm} \epsilon_\mathfrak{b} \pi_\mathfrak{b}^{2k+1} \partial(\pi_\mathfrak{b}^{2k} \xi_\mathfrak{b}) \Psi_{|\zeta|^{\frac{1}{2}}\gamma} d\mu_\mathfrak{b}$$

for $\gamma \in C_c^\infty(U)$.

This follows immediately from Lemma 8 if we observe that $\eta_\mathfrak{a} = \pi_\mathfrak{a}^2$, $\eta_\mathfrak{b} = \pi_\mathfrak{b}^2$ and $\tau \mid \zeta_\mathfrak{a} \mid^{\frac{1}{2}} = \mid \pi_\mathfrak{a} \mid = \epsilon_\mathfrak{a} \pi_\mathfrak{a}$, $\mid \lambda \mid \mid \zeta_\mathfrak{b} \mid^{\frac{1}{2}} = \mid \pi_\mathfrak{b} \mid = \epsilon_\mathfrak{b} \pi_\mathfrak{b}$ on $\mathfrak{a}_1 \cap U$ and $\mathfrak{b}_0' \cap U$, respectively.

3. Proof of Theorem 1. Since $\mathfrak{h} = \sigma + CH_{\alpha_0}$, it is clear that every polynomial in $S(\mathfrak{h})$ can be written uniquely in the form $p = \sum_{k \geqq 0} p_k \alpha_0^k$ where $p_k \in S(\sigma)$. Let s_0 denote the Weyl reflexion in $\mathfrak{h}$ corresponding to α_0. Then since H_{α_0} is orthogonal to σ, it is obvious that $p_k^{s_0} = p_k$ and $s_0 \alpha_0 = -\alpha_0$. Therefore if $p^{s_0} = p$, we can conclude that $p_k = 0$ if k is odd. Now let ξ be a homogeneous element in $I(\mathfrak{g})$ of positive degree and r an integer $\geqq 0$. Then if $\bar{\xi}$ is the restriction of ξ on $\mathfrak{h}$, it follows that (see [4(d), Lemma 9]) that

$$\bar{\xi} \pi^{2r} = \sum_{k \geqq 0} \alpha_0^{2r+2k} q_k,$$

where $q_k \in S(\sigma)$. Similarly,

$$\pi^{2r+1} = \sum_{m \geqq 0} \alpha_0^{2r+1+2m} q_m',$$

where again $q_m' \in S(\sigma)$. Put

$$p_k = \sum_{m \geqq 0} \langle \alpha_0^{2r+1+2m}, \alpha_0^{2r+1+2m} \rangle \partial(q_{m+1+k}) q_m' \qquad (k \geqq 0)$$

in the notation of [4(d), § 2]. Let $d^0(p)$ denote the degree of any polynomial $p \in S(\mathfrak{g})$. Then it is obvious that q_k and q_m' are homogeneous and

$$d^0(q_k) = d^0(\xi) + 2rd^0(\pi) - (2r + 2k),$$

$$d^0(q_m') = (2r + 1) d^0(\pi) - (2r + 1 + 2m).$$

Therefore p_k is also homogeneous and

$$d^0(p_k) = d^0(\pi) - d^0(\xi) + 2k + 1 \qquad (k \geqq 0).$$

LEMMA 9. *For any given integer $r \geqq 0$, we can choose a homogeneous element $\xi \in I(\mathfrak{g})$ of positive degree such that the corresponding polynomial $p_k \neq 0$ for some k.*

Put $\epsilon = 1$ or $(-1)^{\frac{1}{2}}$ according as α_0 lies in Q_+ or in P_0. Then $\alpha^* = \epsilon\alpha_0$ takes only real values on $\mathfrak{h}_0$. Let $\mathfrak{h}_+$ denote the set of all $H \in \mathfrak{h}_0$ where $\alpha^*(H) > 0$. Similarly, let $\mathfrak{h}_-$ be the set of those $H \in \mathfrak{h}_0$ where $\alpha^*(H) < 0$. We call $\mathfrak{h}_+$ and $\mathfrak{h}_-$, respectively, the upper and lower halves of $\mathfrak{h}_0$.

Let $d\mu$ denote the Euclidean measure on $\mathfrak{h}_0$ and $\mathcal{E}$ the distribution on $\mathfrak{h}_0$ given by

$$\mathcal{E}(g) = \int_{\mathfrak{h}_+} g\, d\mu \qquad (g \in C_c^\infty(\mathfrak{h}_0)).$$

Put $\mathcal{E}_r = \partial(\pi^{2r})(\pi^{2r+1}\mathcal{E})$. We shall first prove the following result.

LEMMA 10. *There exists a homogenous element $\bar{\xi} \in I(\mathfrak{h})$ of positive degree such that H_0 lies in the carrier of $\partial(\bar{\xi})\mathcal{E}_r$.*

Let $\square$ denote the polynomial function on $\mathfrak{h}$ defined by $\square(H) = -B(H, \theta(H))$ $(H \in \mathfrak{h})$. Let W be the Weyl group of $\mathfrak{g}$ with respect to $\mathfrak{h}$, and consider the polynomial

$$\prod_{s \in W} (t - \square^s) = t^w + q_1 t^{w-1} + \cdots + q_w \qquad (q_i \in S(\mathfrak{h}))$$

in the indeterminate t with coefficients in $S(\mathfrak{h})$. It is obvious that q_i $(1 \leqq i \leqq w)$ are homogeneous elements in $I(\mathfrak{h})$ of positive degree. Replacing t by $\square$ we get,

$$\square^w + \square^{w-1} q_1 + \cdots + q_w = \prod_{s \in W} (\square - \square^s) = 0.$$

Now if our assertion is false, we can choose an open connected neighborhood V of H_0 in $\mathfrak{h}_0$ such that $\partial(q_i)\mathcal{E}_r = 0$ on V $(1 \leqq i \leqq w)$. Therefore since

$$\partial(\square^w) + \partial(\square^{w-1})\partial(q_1) + \cdots + \partial(q_w) = 0,$$

we conclude that $\partial(\square^w)\mathcal{E}_r = 0$ on V. But $\square$ is obviously a positive-definite quadratic form in $\mathfrak{h}_0$, and therefore this differential equation is of the *elliptic type* (see Garding [3]). Hence it follows (see [7, p. 136] and [6]) that $\mathcal{E}_r$ must coincide on V with an analytic function. But $\mathcal{E} = 0$ on $\mathfrak{h}_-$, and therefore $\mathcal{E}_r = 0$ on $V \cap \mathfrak{h}_-$. Moreover, $V \cap \mathfrak{h}_-$ is open and not empty since $\alpha_0(H_0) = 0$. Therefore we conclude that $\mathcal{E}_r = 0$ on V. On the other hand, $V \cap \mathfrak{h}_+$ is also not empty and $\mathcal{E} = 1$ on $V \cap \mathfrak{h}_+$. Therefore $\partial(\pi)\mathcal{E}_r = \partial(\pi^{2r+1})\pi^{2r+1} = \langle \pi^{2r+1}, \pi^{2r+1} \rangle$ on $V \cap \mathfrak{h}_+$. But we know from the Corollary to Lemma 18 of [4(d)] that $\langle \pi^{2r+1}, \pi^{2r+1} \rangle$ is a positive constant. This contradiction proves the lemma.

Now in order to prove Lemma 9, we select (see Lemma 9 of [4(d)]) a homogeneous element $\xi \in I(\mathfrak{g})$ of positive degree such that its restriction $\bar{\xi}$

on $\mathfrak{h}$ fulfills the condition of Lemma 10. It is obvious that $\partial(H)\mathcal{E}=0$ for $H\in\sigma$. Moreover,

$$\mathcal{E}(\partial(\alpha_0)(\alpha_0 g))=\int_{\mathfrak{h}_+}\partial(\alpha_0)(\alpha_0 g)\,d\mu=0 \qquad (g\in C_c^\infty(\mathfrak{h}_0)),$$

and therefore $\alpha_0\partial(\alpha_0)\mathcal{E}=0$. We claim that

$$\partial(\alpha_0^k)(\alpha_0^m\mathcal{E})=(\partial(\alpha_0^k)\alpha_0^m)\mathcal{E} \qquad (m\geqq k\geqq 0).$$

This is obviously true for $k=0$. So we assume $k\geqq 1$ and use induction on k. Then $\partial(\alpha_0^{k-1})(\alpha_0^m\mathcal{E})=(\partial(\alpha_0^{k-1})\alpha_0^m)\mathcal{E}$, and therefore

$$\partial(\alpha_0^k)(\alpha_0^m\mathcal{E})=(\partial(\alpha_0^k)\alpha_0^m)\mathcal{E}+(\partial(\alpha_0^{k-1})\alpha_0^m)\partial(\alpha_0)\mathcal{E}=(\partial(\alpha_0^k)\alpha_0^m)\mathcal{E}$$

since $m>k-1$ and $\alpha_0\partial(\alpha_0)\mathcal{E}=0$. Hence if $k>m$,

$$\partial(\alpha_0^k)(\alpha_0^m\mathcal{E})=\partial(\alpha_0^{k-m})\{(\partial(\alpha_0^m)\alpha_0^m)\mathcal{E}\}=\langle\alpha_0^m,\alpha_0^m\rangle\partial(\alpha_0^{k-m})\mathcal{E}.$$

Now consider $\partial(\bar{\xi})\mathcal{E}_r$. Then if

$$\bar{\xi}\pi^{2r}=\sum_{k\geqq 0}\alpha_0^{2r+2k}q_k,\qquad \pi^{2r+1}=\sum_{m\geqq 0}\alpha_0^{2r+1+2m}q_m'$$

$(q_k, q_m'\in S(\sigma))$ as before, it is clear that

$$\begin{aligned}\partial(\bar{\xi})\mathcal{E}_r=\partial(\bar{\xi}\pi^{2r})(\pi^{2r+1}\mathcal{E})&=\sum_{k,m\geqq 0}\partial(\alpha_0^{2r+2k}q_k)(\alpha_0^{2r+1+2m}q_m'\mathcal{E})\\&=\sum_{k,m\geqq 0}(\partial(q_k)q_m')\partial(\alpha_0^{2r+2k})(\alpha_0^{2r+1+2m}\mathcal{E})\\&=\sum_{m\geqq k\geqq 0}(\partial(q_k)q_m')(\partial(\alpha_0^{2r+2k})\alpha_0^{2r+1+2m})\mathcal{E}\\&\quad+\sum_{k>m\geqq 0}(\partial(q_k)q_m')\langle\alpha_0^{2r+1+2m},\alpha_0^{2r+1+2m}\rangle\partial(\alpha_0^{2k-2m-1})\mathcal{E}\end{aligned}$$

since σ and H_{α_0} are orthogonal. Now the first sum is obviously equal to $(\partial(\bar{\xi}\pi^{2r})\pi^{2r+1})\mathcal{E}$. But $\partial(\bar{\xi}\pi^{2r})\pi^{2r+1}$ is a homogeneous polynomial in $S(\mathfrak{h})$ of degree $d^0(\pi)-d^0(\xi)$ which is skew-invariant under W. Therefore it follows from Lemma 10 of [4(d)] that it is zero. Put $\delta^{(k)}=\partial(\alpha_0^{k+1})\mathcal{E}$ $(k\geqq 0)$. Then

$$\partial(\bar{\xi})\mathcal{E}_r=\sum_{k\geqq 0}p_k\delta^{(2k)},$$

where the polynomials $p_k\in S(\sigma)$ are defined as before by

$$p_k=\sum_{m\geqq 0}(\partial(q_{m+1+k})q_m')\langle\alpha_0^{2r+1+2m},\alpha_0^{2r+1+2m}\rangle.$$

On the other hand, since $\partial(\bar{\xi})\mathcal{E}_r\neq 0$, not every p_k can be zero. This completes the proof of Lemma 9.

Now fix r and ξ once for all as in Lemma 9 and let k_0 be an integer such that $p_{k_0}\neq 0$. As before, let $d\mu_\sigma$ denote the Euclidean measure on σ_0.

LEMMA 11. *If U_σ is any open neighborhood of H_0 in σ_0, we can select a function $f \in C_c^\infty(U_\sigma)$ such that*

$$\int_{U_\sigma} f p_k \, d\mu_\sigma = \begin{cases} 0 \text{ if } k \neq k_0, \\ 1 \text{ if } k = k_0 \end{cases} \qquad (k \geqq 0).$$

Let $k_1 < k_2 < \cdots < k_m$ be all the values of k for which $p_k \neq 0$. Then as we have seen p_{k_i} is a homogeneous polynomial in $S(\sigma)$ of degree $d^0(\pi) - d^0(\xi) + 2k_i + 1$ $(1 \leqq i \leqq m)$. Therefore, regarded as functions on $U_\sigma, p_{k_1}, \cdots, p_{k_m}$ are linearly independent over C, and our assertion follows from Lemma 20 of the Appendix (Section 6).

Now let $\mathfrak{a}_1 = \sigma_0 + R^+H'$ and $\mathfrak{b}_0' = \sigma + R'(X' - Y')$ as before. We define the $\mathcal{B}$-distributions (see [4(d), § 2]) $\mathcal{E}_\mathfrak{a}$ and $\mathcal{E}_\mathfrak{b}^\pm$ on $\mathfrak{a}_1$ and $\mathfrak{b}_0'$, respectively, as follows:

$$\mathcal{E}_\mathfrak{a}(g_1) = \int_{\mathfrak{a}_1} g_1 \, d\mu_\mathfrak{a} \qquad (g_1 \in \mathcal{B}(\mathfrak{a}_1)),$$

$$\mathcal{E}_\mathfrak{b}^+(g_2) = \int_{\mathfrak{b}_2} g_2 \, d\mu_\mathfrak{b}, \qquad \mathcal{E}_\mathfrak{b}^-(g_2) = \int_{\mathfrak{b}_3} g_2 \, d\mu_\mathfrak{b} \qquad (g_2 \in \mathcal{B}(\mathfrak{b}_0')).$$

Let $\delta_\mathfrak{a}^{(k)} = \partial(\tau^{k+1})\mathcal{E}_\mathfrak{a}$ and ${}^\pm\delta_\mathfrak{b}^{(k)} = \partial(\lambda^{k+1})\mathcal{E}_\mathfrak{b}^\pm$ $(k \geqq 0)$. For any automorphism y of $\mathfrak{g}$ and $p \in S(\mathfrak{g})$, let p^y denote the polynomial function $X \to p(y^{-1}X)$ $(X \in \mathfrak{g})$. Then it is clear that $\pi_\mathfrak{a} = \pi^{\nu_\mathfrak{a}}$, $\pi_\mathfrak{b} = \pi^{\nu_\mathfrak{b}}$ and $p^{\nu_\mathfrak{a}} = p^{\nu_\mathfrak{b}} = p$ for $p \in S(\sigma)$. Moreover, $(\alpha_0)^{\nu_\mathfrak{a}} = \tau$, $(\alpha_0)^{\nu_\mathfrak{b}} = \lambda$ and $(\bar{\xi})^{\nu_\mathfrak{a}} = \xi_\mathfrak{a}$, $(\bar{\xi})^{\nu_\mathfrak{b}} = \xi_\mathfrak{b}$ in the notation of Lemma 8. Therefore

$$\xi_\mathfrak{a}\pi_\mathfrak{a}^{2r} = \sum_{k \geqq 0} \tau^{2r+2k} q_k, \qquad \pi_\mathfrak{a}^{2r+1} = \sum_{m \geqq 0} \tau^{2r+1+2m} q_m'$$

$$\xi_\mathfrak{b}\pi_\mathfrak{b}^{2r} = \sum_{k \geqq 0} \lambda^{2r+2k} q_k, \qquad \pi_\mathfrak{b}^{2r+1} = \sum_{m \geqq 0} \lambda^{2r+1+2m} q_m'.$$

Moreover, it again follows without difficulty that $\partial(H)\mathcal{E}_\mathfrak{a} = 0$, $\partial(H)\mathcal{E}_\mathfrak{b}^\pm = 0$ for $H \in \sigma$ and $\tau\partial(\tau)\mathcal{E}_\mathfrak{a} = 0$, $\lambda\partial(\lambda)\mathcal{E}_\mathfrak{b}^\pm = 0$. Therefore we conclude in the same way as before, that

$$\partial(\xi_\mathfrak{a}\pi_\mathfrak{a}^{2r})(\pi_\mathfrak{a}^{2r+1}\mathcal{E}_\mathfrak{a}) = \sum_{k \geqq 0} p_k \delta_\mathfrak{a}^{(2k)}$$

$$\partial(\xi_\mathfrak{b}\pi_\mathfrak{b}^{2r})(\pi_\mathfrak{b}^{2r+1}\mathcal{E}_\mathfrak{b}) = \sum_{k \geqq 0} p_k {}^\pm\delta_\mathfrak{b}^{(2k)}.$$

Now U being defined as in Lemma 8, let U_σ and $U_\mathfrak{l}$ denote the orthogonal projections of U in σ_0 and $\mathfrak{l}_0$, respectively. Select $g_0 \in C_c^\infty(U_\sigma)$ according to Lemma 11, and for any $\beta \in C_c^\infty(U_\mathfrak{l})$, define $\gamma_\beta \in C_c^\infty(U)$ by

$$\gamma_\beta(H_\sigma + Z) = g_0(H_\sigma)\beta(Z)\,|\,\zeta(H_\sigma + Z)\,|^{-\frac{1}{2}} \qquad (H_\sigma \in U_\sigma, Z \in U_\mathfrak{l}).$$

Moreover, let $F_\beta = f_{\gamma_\beta}$ (in the notation of Lemmas 1 and 8) and

$$\phi_\beta(t) = \int_{-\infty}^{\infty} \bar{\beta}(tH' + xX')\,dx \qquad (t \in R^+)$$

$$\psi_\beta(s) = |s| \int_0^{\infty} \bar{\beta}(s(e^t X' - e^{-t} Y'))(e^t - e^{-t})\,dt \qquad (s \in R'),$$

where[2] $\bar{\beta}(Z) = \pi^{-1} \int_0^{\pi} \beta(u_t Z)\,dt$ and $u_t = \exp(t\,\mathrm{ad}(X' - Y'))$. Put $\phi_\beta(0) = \lim_{t \to 0} \phi_\beta(t)$ $(t \in R^+)$, $\psi_\beta^+(0) = \lim_{s \to 0} \psi_\beta(s)$ $(s \in R^+)$ and $\psi_\beta^-(0) = \lim_{s \to 0} \psi_\beta(s)$ $(s \in R^-)$. All these limits exist (see Section 6). Also let π' denote the product of all roots $\alpha \neq \alpha_0$ in P. Then $\pi'(H_0) \neq 0$. We set $\epsilon = |\pi'(H_0)|/\pi'(H_0)$.

LEMMA 12. *Let $q = d^0(\xi)$ and $a = \epsilon(-1)^q \{\alpha_0(H_{\mathfrak{a}_0})\}^{2k_0+1} 2^{k_0-2}$, where k_0 has the same meaning as in Lemma 11. Then for any $\beta \in C_c^\infty(U_{\mathfrak{l}})$,*

$$\int_{N_1} \eta^r \partial(\eta^r \xi) F_\beta\,dX = a\phi_{\beta_0}(0),$$

where $\beta_0 = \partial(\Omega^{k_0})\beta$. Similarly,

$$\int_{N^\pm} \eta^r \partial(\eta^r \xi) F_\beta\,dX = -a\psi_{\beta_0}^{\pm}(0)$$

if $[\Xi : LB] = 1$, and, if $[\Xi : LB] = 2$,

$$\int_{N_2 \cup N_3} \eta^r \partial(\eta^r \xi) F_\beta\,dX = -a\{\psi_{\beta_0}^{+}(0) + \psi_{\beta_0}^{-}(0)\}.$$

Define $\epsilon_{\mathfrak{a}}$ and $\epsilon_{\mathfrak{b}}$ on $\mathfrak{a}_1 \cap U$ and $\mathfrak{b}_0' \cap U$ as in the Corollary to Lemma 8. Then $\epsilon_{\mathfrak{a}}$ is constant on $\mathfrak{a}_1 \cap U$ and $\epsilon_{\mathfrak{b}}$ on $\mathfrak{b}_i \cap U$ $(i = 2, 3)$, and it is obvious that $\pi_{\mathfrak{a}}'(H_0) = \pi_{\mathfrak{b}}'(H_0) = \pi'(H_0) \neq 0$, where $\pi_{\mathfrak{a}}' = (\pi')^{\nu_{\mathfrak{a}}}$, $\pi_{\mathfrak{b}}' = (\pi')^{\nu_{\mathfrak{b}}}$. Therefore since $\pi_{\mathfrak{a}} = \tau\pi_{\mathfrak{a}}'$, by making H tend to H_0 $(H \in \mathfrak{a}_1 \cap U)$, we conclude immediately that $\epsilon_{\mathfrak{a}} = \epsilon$. Similarly, $\epsilon_{\mathfrak{b}} = -(-1)^{\frac{1}{2}}\epsilon$ and $(-1)^{\frac{1}{2}}\epsilon$ on $\mathfrak{b}_2$ and $\mathfrak{b}_3$, respectively. Hence it follows from the Corollary to Lemma 8 that

$$\int_{N_1} \eta^r \partial(\eta^r \xi) F_\beta\,dX = \tfrac{1}{2}\epsilon\mathcal{E}_{\mathfrak{a}}(\pi_{\mathfrak{a}}^{2r+1} \partial(\pi_{\mathfrak{a}}^{2r}\xi_{\mathfrak{a}}) \Phi_{|\zeta|^{\frac{1}{2}}\gamma_\beta})$$
$$= \tfrac{1}{2}\epsilon(-1)^q \sum_{k \geqq 0} \delta_{\mathfrak{a}}^{(2k)}(p_k \Phi_{|\zeta|^{\frac{1}{2}}\gamma_\beta}) = -\tfrac{1}{2}\epsilon(-1)^q \sum_{k \geqq 0} \mathcal{E}_{\mathfrak{a}}(p_k \partial(\tau^{2k+1}) \Phi_{|\zeta|^{\frac{1}{2}}\gamma_\beta})$$

in view of our result above. But it is clear that

$$\Phi_{|\zeta|^{\frac{1}{2}}\gamma_\beta}(H_\sigma + tH') = g_0(H_\sigma)\psi_\beta(t) \qquad (H_\sigma \in U_\sigma,\ t \in R^+).$$

Let H_τ denote the element in $\mathfrak{a}$ such that $B(H, H_\tau) = \tau(H)$ for all $H \in \mathfrak{a}$. Then it is clear that $\tau(H_\tau) = \alpha_0(H_{\mathfrak{a}_0}) = \langle \alpha_0, \alpha_0 \rangle$, and hence $H_\tau = \tfrac{1}{2}\langle \alpha_0, \alpha_0 \rangle H'$. Therefore

$$\mathcal{E}_a(p_k\partial(\tau^{2k+1})\Phi_{|\zeta|^{\frac{1}{2}}\gamma_\beta})$$
$$=\langle\alpha_0,\alpha_0\rangle^{2k+1}2^{-(2k+1)}\int_{U\sigma}p_k g_0 d\mu_\sigma\int_0^\infty(d^{2k+1}\phi_\beta/dt^{2k+1})\,dt.$$

On the other hand, we know from Lemma 17 (Appendix, Section 6) that

$$\phi_{\partial(\Omega^k)\beta}=2^{-3k}d^{2k}\phi_\beta/dt^{2k}.$$

So it follows from the definition of g_0 that

$$\int_{N_1}\eta^r\partial(\eta^r\xi)F_\beta\,dX=\epsilon\langle\alpha_0,\alpha_0\rangle^{2k_0+1}2^{k_0-2}(-1)^q\phi_{\beta_0}(0),$$

and this proves the first formula.

Now suppose $[\Xi\colon LB]=1$. Then it follows again from the Corollary to Lemma 8 that

$$\int_{N^\pm}\eta^r\partial(\eta^r\xi)F_\beta dX=\mp\tfrac{1}{2}\epsilon(-1)^{\frac{1}{2}}\mathcal{E}_{\mathfrak{b}}{}^{\pm}(\pi_{\mathfrak{b}}{}^{2r+1}\partial(\pi_{\mathfrak{b}}{}^{2r}\xi_{\mathfrak{b}})\Psi_{|\zeta|^{\frac{1}{2}}\gamma_\beta})$$
$$=\mp\tfrac{1}{2}\epsilon(-1)^{q+\frac{1}{2}}\sum_{k\geqq 0}{}^{\pm}\delta_{\mathfrak{b}}{}^{(2k)}(p_k\Psi_{|\zeta|^{\frac{1}{2}}\gamma_\beta})=\pm\tfrac{1}{2}\epsilon(-1)^{q+\frac{1}{2}}\sum_{k\geqq 0}\mathcal{E}_{\mathfrak{b}}{}^{\pm}(p_k\partial(\lambda^{2k+1})\Psi_{|\zeta|^{\frac{1}{2}}\gamma_\beta}).$$

Let H_λ denote the element in $\mathfrak{b}$ such that $B(H,H_\lambda)=\lambda(H)$ $(H\in\mathfrak{b})$. Then $\lambda(H_\lambda)=\alpha_0(H_{\alpha_0})=\langle\alpha_0,\alpha_0\rangle$ and $H_\lambda=\nu_{\mathfrak{b}}H_{\alpha_0}=-\frac{1}{2}(-1)^{\frac{1}{2}}\langle\alpha_0,\alpha_0\rangle(X'-Y')$. Therefore

$$\mathcal{E}_{\mathfrak{b}}{}^{+}(p_k\partial(\lambda^{2k+1})\Psi_{|\zeta|^{\frac{1}{2}}\gamma_\beta})$$
$$=\langle\alpha_0,\alpha_0\rangle^{2k+1}(-1)^{k-\frac{1}{2}}2^{-(2k+1)}\int_{U\sigma}p_k g_0 d\mu_\sigma\int_0^\infty(d^{2k+1}\psi_\beta/dt^{2k+1})\,dt$$

and

$$\mathcal{E}_{\mathfrak{b}}{}^{-}(p_k\partial(\lambda^{2k+1})\Psi_{|\zeta|^{\frac{1}{2}}\gamma_\beta})$$
$$=\langle\alpha_0,\alpha_0\rangle^{2k+1}(-1)^{k-\frac{1}{2}}2^{-(2k+1)}\int_{U\sigma}p_k g_0 d\mu_\sigma\int_{-\infty}^0(d^{2k+1}\psi_\beta/dt^{2k+1})\,dt.$$

On the other hand,

$$\psi_{\partial(\Omega^k)\beta}=(-1)^k 2^{-3k}d^{2k}\psi_\beta/dt^{2k}$$

from Lemma 17 (Section 6). Therefore

$$\mathcal{E}_{\mathfrak{b}}{}^{\pm}(p_k\partial(\lambda^{2k+1})\Psi_{|\zeta|^{\frac{1}{2}}\gamma_\beta})$$
$$=\pm\langle\alpha_0,\alpha_0\rangle^{2k+1}(-1)^{\frac{1}{2}}2^{k-1}\psi_{\partial(\Omega^k)\beta}{}^{\pm}(0)\int_{U\sigma}p_k g_0 d\mu_\sigma,$$

and by taking into account the definition of g_0 we get,

$$\int_{N^\pm}\eta^r\partial(\eta^r\xi)F_\beta\,dX=-\epsilon\langle\alpha_0,\alpha_0\rangle^{2k_0+1}(-1)^q 2^{k_0-2}\psi_{\beta_0}{}^{\pm}(0).$$

Finally, suppose $[\Xi : LB] = 2$. Then again from the Corollary to Lemma 8,

$$\int_{N_2 \cup N_3} \eta^r \partial(\eta^r \xi) F_\beta \, dX = -\tfrac{1}{2}\epsilon(-1)^{\frac{1}{2}} \mathcal{E}_{\mathfrak{b}}(\pi_{\mathfrak{b}}{}^{2r+1} \partial(\pi_{\mathfrak{b}}{}^{2r} \xi_{\mathfrak{b}}) \Psi_{|\zeta|^{\frac{1}{2}}\gamma_\beta}),$$

where $\mathcal{E}_{\mathfrak{b}} = \mathcal{E}_{\mathfrak{b}}{}^+ - \mathcal{E}_{\mathfrak{b}}{}^-$, and we conclude from the above caluculation that

$$\int_{N_2 \cup N_3} \eta^r \partial(\eta^r \xi) F_\beta \, dX = -\epsilon \langle \alpha_0, \alpha_0 \rangle^{2k_0+1} (-1)^q 2^{k_0-2} [\psi_{\beta_0}{}^+(0) + \psi_{\beta_0}{}^-(0)].$$

This completes the proof of Lemma 12.

We are now in a position to prove the main result of this section.

THEOREM 1. *Let r and ξ be defined as in Lemma 9 and let c_1, c_2, c_3 be three complex numbers. We assume that $c_2 = c_3$ in case $[\Xi : LB] = 2$. Let N_0 be an open neighborhood of H_0 in N and let T be the distribution on N_0 defined as follows:*

$$T(f) = \sum_{1 \leq i \leq 3} c_i \int_{N_i} \eta^r \partial(\eta^r) f dX \qquad (f \in C_c^\infty(N_0)).$$

Then $\partial(\xi) T = 0$ implies that $c_1 = c_2 = c_3$.

Let ${}_0U$ be an open neighborhood of H_0 in U and ${}_0V^*$ an open neighborhood of 1^* in V^* (see Lemma 1 for the definition of V^*). We can choose ${}_0V^*$ and ${}_0U$ so small that $\phi(x^*)Z \in N_0$ for $x^* \in {}_0V^*$, $Z \in {}_0U$. Let ${}_0U_\sigma$ and ${}_0U_{\mathrm{I}}$ be the orthogonal projections of ${}_0U$ in σ_0 and I_0, respectively. We may further assume that ${}_0U = {}_0U_\sigma + {}_0U_{\mathrm{I}}$. Moreover, we can suppose that the carriers of the functions a_0 and g_0 (used in the definitions of f_γ and F_β in Lemmas 1 and 12, respectively) are contained in ${}_0V^*$ and ${}_0U_\sigma$, respectively. Then it is clear that $F_\beta \in C_c^\infty(N_0)$ for any $\beta \in C_c^\infty({}_0U_{\mathrm{I}})$, and therefore $T(\partial(\xi)F_\beta) = 0$ by our hypothesis. On the other hand, it follows from Lemma 12 that

$$T(\partial(\xi)F_\beta) = a\{c_1 \phi_{\partial(\Omega^{k_0})\beta}(0) - c_2 \psi^+{}_{\partial(\Omega^{k_0})\beta}(0) - c_3 \psi^-{}_{\partial(\Omega^{k_0})\beta}(0)\}$$

since $c_2 = c_3$ when $[\Xi : LB] = 2$. Now put $\beta = \Omega^{k_0}\gamma$ $(\gamma \in C_c^\infty({}_0U_{\mathrm{I}}))$ and $\gamma_0 = \partial(\Omega^{k_0})(\Omega^{k_0}\gamma)$. Then it is clear that

$$c_1 \phi_{\gamma_0}(0) - c_2 \psi_{\gamma_0}{}^+(0) - c_3 \psi_{\gamma_0}{}^-(0) = 0$$

for all $\gamma \in C_c^\infty({}_0U_{\mathrm{I}})$. On the other hand if $c = 2^{2k_0} k_0! \Gamma(k_0 + \frac{1}{2}) / \Gamma(\frac{1}{2})$, it follows from the Corollary of Lemma 18 (Section 6) that $\phi_{\gamma_0}(0) = c\phi_\gamma(0)$, $\psi_{\gamma_0}{}^\pm(0) = c\psi_\gamma{}^\pm(0)$. Therefore, since c is positive, we conclude that

$$c_1 \phi_\gamma(0) - c_2 \psi_\gamma{}^+(0) - c_3 \psi_\gamma{}^-(0) = 0$$

for all $\gamma \in C_c^\infty({}_0U_{\mathrm{I}})$, and now it follows from Lemma 19 (Section 6) that $c_1 = c_2 = c_3$.

4. Proof of Theorem 2. We shall now apply Theorem 1 to prove that the constants $c_1, \cdots, c_N$ of Lemma 29 of [4(e)] are all equal. But first we need one additional fact.

LEMMA 13. *Let S denote the set of those elements in $\mathfrak{g}_0$ which are either regular or semiregular. Then S is connected.*

Suppose V_1, V_2 are two closed subsets of $\mathfrak{g}_0$ such that $V_1 \cup V_2 = \mathfrak{g}_0$ and $V_1 \cap V_2 \cap S = \emptyset$. Then we have to show that S is contained in one of the two sets V_1, V_2. Let $\mathfrak{h}_0$ be a Cartan subalgebra of $\mathfrak{g}_0$ such that $\theta(\mathfrak{h}_0) = \mathfrak{h}_0$. We use the notation of Section 2. For any root $\alpha \in P$, let σ_α denote the set of those $H \in \mathfrak{h}_0$ where $\alpha(H) = 0$. Moreover, let U be the complement in $\mathfrak{h}_0$ of the union of the sets $\sigma_\alpha \cap \sigma_\beta$ ($\alpha, \beta \in P$; $\alpha \neq \beta$). Then U is connected (see the Corollary to Lemma 42 of [4(e)]), and it is obvious that $U \subset S$. Therefore U is contained in one of the two sets V_1, V_2. But since U is dense in $\mathfrak{h}_0$, we conclude that either V_1 or V_2 contains $\mathfrak{h}_0$.

Let X be an element in S. Then since G is connected, $\bigcup_{x \in G} xX$ is a connected subset of S. Then if $X \in V_i$, it follows that $xX \in V_i$ for every $x \in G$ $(i = 1, 2)$. Now select i such that $\mathfrak{h}_0 \subset V_i$. Then $x(\mathfrak{h}_0 \cap \mathfrak{g}_0') \subset V_i$ $(x \in G)$, and therefore since $\mathfrak{h}_0 \cap \mathfrak{g}_0'$ is dense in $\mathfrak{h}_0$, $x\mathfrak{h}_0 \subset V_i$.

Let $\mathfrak{a}_0$ be a fundamental Cartan subalgebra (see [4(e), § 8]) of $\mathfrak{g}_0$ such that $\mathfrak{a}_0 = \theta(\mathfrak{a}_0)$. We may assume that $\mathfrak{a}_0 \subset V_1$. For any Cartan subalgebra $\mathfrak{r}_0$ of $\mathfrak{g}_0$, define $l_-(\mathfrak{r}_0)$ as in [4(e), § 8]. We shall prove by induction on $l_-(\mathfrak{a}_0) - l_-(\mathfrak{h}_0)$ that $\mathfrak{h}_0 \subset V_1$. If $l_-(\mathfrak{a}_0) = l_-(\mathfrak{h}_0)$, $\mathfrak{h}_0$ is also fundamental and therefore conjugate to $\mathfrak{a}_0$ under G (see the Corollary to Lemma 32 of [4(e)]). Hence it follows from what has been said above that $\mathfrak{h}_0 \subset V_1$. So let us now suppose that $l_-(\mathfrak{h}_0) < l_-(\mathfrak{a}_0)$. Since K is compact and $\mathfrak{a}_0 \cap \mathfrak{k}_0$ is a maximal abelian subalgebra of $\mathfrak{k}_0$ (Corollary to Lemma 31 of [4(e)]), we may assume (by replacing $\mathfrak{h}_0$ by a conjugate Cartan subalgebra) that $\mathfrak{h}_0 \cap \mathfrak{k}_0 \subset \mathfrak{a}_0 \cap \mathfrak{k}_0$. Let $\mathfrak{q}_0$ be the centralizer of $\mathfrak{h}_0 \cap \mathfrak{k}_0$ in $\mathfrak{g}_0$. Then obviously $\theta(\mathfrak{q}_0) = \mathfrak{q}_0$ and therefore $\mathfrak{q}_0$ is reductive in $\mathfrak{g}_0$ (Lemma 10 of [4(e)]). Moreover, since $\mathfrak{h}_0$ is maximal abelian in $\mathfrak{g}_0$, it is clear that $\mathfrak{h}_0 \cap \mathfrak{p}_0$ is a maximal abelian subspace of $\mathfrak{q}_0 \cap \mathfrak{p}_0$. Let L be the connected component of 1 in the centralizer of $\mathfrak{h}_0 \cap \mathfrak{k}_0$ in K. Then L is compact and its Lie algebra is $\mathfrak{q}_0 \cap \mathfrak{k}_0$. Moreover, since $\mathfrak{q}_0$ is reductive and $\theta(\mathfrak{q}_0) = \mathfrak{q}_0$, any two maximal abelian subspaces of $\mathfrak{q}_0 \cap \mathfrak{p}_0$ are conjugate under L (see Cartan [1] or Lemma 39 of

[4(b)]). Therefore since $\mathfrak{a}_0 \cap \mathfrak{p}_0 \subset \mathfrak{q}_0 \cap \mathfrak{p}_0$, we can select $k \in L$ such that $\mathfrak{a}_0 \cap \mathfrak{p}_0 \subset k(\mathfrak{h}_0 \cap \mathfrak{p}_0)$. Then by replacing $\mathfrak{h}_0$ by $k\mathfrak{h}_0$ we can assume that $\mathfrak{h}_0 \cap \mathfrak{p}_0 \supset \mathfrak{a}_0 \cap \mathfrak{p}_0$ and $\mathfrak{h}_0 \cap \mathfrak{k}_0 \subset \mathfrak{a}_0 \cap \mathfrak{k}_0$. Now since $\mathfrak{h}_0$ is not fundamental, there exists a root $\alpha_0 \in P$ which takes only real values on $\mathfrak{h}_0$ (see Lemma 33 of [4(e)]), and we can assume that $X_{\alpha_0}, X_{-\alpha_0} \in \mathfrak{g}_0$ and $X_{\alpha_0} - X_{-\alpha_0} \in \mathfrak{k}_0$ (Lemma 46 of [4(e)]). Then $\mathfrak{h}_1 = \sigma_{\alpha_0} + R(X_{\alpha_0} - X_{-\alpha_0})$ is a Cartan subalgebra of $\mathfrak{g}_0$ and $\theta(\mathfrak{h}_1) = \mathfrak{h}_1$. Since α_0 is identically zero on $\mathfrak{h}_0 \cap \mathfrak{k}_0$, it follows that $\sigma_{\alpha_0} \supset \mathfrak{h}_0 \cap \mathfrak{k}_0$, and therefore $l_-(\mathfrak{h}_1) \geqq l_-(\mathfrak{h}_0) + 1$. Hence $\mathfrak{h}_1 \subset V_1$ by our induction hypothesis. On the other hand, we can obviously choose a point $H \in \sigma_{\alpha_0}$ such that $\alpha(H) \neq 0$ for any root $\alpha \neq \alpha_0$ in P. Then H is semiregular and therefore $\mathfrak{h}_0 \cap \mathfrak{h}_1 \cap S \neq \emptyset$. But since $V_1 \cap V_2 \cap S = \emptyset$, we conclude that $\mathfrak{h}_0$ must also be contained in V_1.

Now let X be any element in $\mathfrak{g}_0'$ and let $\mathfrak{h}_X$ denote its centralizer in $\mathfrak{g}_0$. Then $\mathfrak{h}_X$ is a Cartan subalgebra of $\mathfrak{g}_0$ and some conjugate of $\mathfrak{h}_X$ (under G) is invariant under θ (see [4(c), Corollary to Lemma 1]). Hence it follows from our result that $\mathfrak{h}_X \subset V_1$. Therefore, since $X \in \mathfrak{h}_X$ and $\mathfrak{g}_0'$ is dense in $\mathfrak{g}_0$, we conclude that $V_1 = \mathfrak{g}_0$. This proves that $S \subset V_1$.

Define $\mathfrak{h}_0$ and T' as in Lemma 27 of [4(e)]. Then T' is a distribution on $\mathfrak{g}_0$.

LEMMA 14. *There exists an open neighborhood V of S in $\mathfrak{g}_0$ and a constant c' such that $T' = c'$ on V.*

Since S is connected, it would be sufficient to prove that T' coincides with a constant on some open neighborhood (in $\mathfrak{g}_0$) of any point $X_0 \in S$. If X_0 is regular, this follows from Lemma 27 of [4(e)]. So now suppose that X_0 is semiregular. Then since $\operatorname{ad} X_0$ is semisimple, X_0 is contained in some Cartan subalgebra $\mathfrak{a}_0$ of $\mathfrak{g}_0$. Select $x \in G$ such that $\theta(x\mathfrak{a}_0) = x\mathfrak{a}_0$ (see the Corollary to Lemma 1 of [4(c)]), and put $H_0 = xX_0$, $x\mathfrak{a}_0 = \mathfrak{b}_0$. Then since T' is invariant under G, it would be sufficient to prove that T' coincides with a constant on some neighborhood of H_0. On the other hand, we know from Lemma 30 of [4(e)] that $\partial(\xi)T' = 0$ for any homogeneous element $\xi \in I(\mathfrak{g})$ of positive degree. Now select a neighborhood N of H_0 corresponding to Lemmas 1 and 2, and let $\mathfrak{g}_i$ $(1 \leqq i \leqq k)$ be all the distinct connected components of $\mathfrak{g}_0'$ which meet N. Then $k \leqq 3$ from Lemma 2. Moreover, it follows from Lemma 29 of [4(e)] that there exists an open neighborhood N_0 of H_0 in N, an integer $r \geqq 0$ and complex numbers c_i $(1 \leqq i \leqq k)$ such that

$$\langle \eta^r, \eta^r \rangle T'(f) = \sum_{1 \leqq i \leqq k} c_i \int_{N \cap \mathfrak{g}_i} \eta^r \partial(\eta^r) f \, dX$$

for all $f \in C_c^\infty(N_0)$. Now the set of singular elements of $\mathfrak{g}_0$ obviously has Euclidean measure zero. Therefore if $k=1$, it is obvious that

$$\langle\eta^r,\eta^r\rangle T'(f) = c_1 \int_{\mathfrak{g}_0} \eta^r\partial(\eta^r) f\, dX = c_1\langle\eta^r,\eta^r\rangle \int_{\mathfrak{g}_0} f\, dX,$$

since η is a homogeneous polynomial of even degree. Moreover, $\langle\eta^r,\eta^r\rangle \neq 0$ (Corollary to Lemma 18 of [4(d)]). Therefore $T'=c_1$ on N_0 in this case. So let us now assume that $k \geqq 2$. Then we may obviously suppose that N is chosen in accordance with Lemma 8. Select ξ corresponding to Lemma 9 and let us use the notation of Theorem 1. If $[\Xi: LB]=2$, we know from Lemma 7 that there exists an element $x \in \Xi$ such that $x(X'-Y') = -(X'-Y')$. On the other hand, if t is positive and sufficiently small, $H_0+t(X'-Y') \in N_2$ while $x(H_0+t(X'-Y')) = H_0-t(X'-Y') \in N_3$. Moreover, since G is connected, the points $y(H_0+t(X'-Y'))$ $(y \in G)$ are all contained in the same connected component of $\mathfrak{g}_0'$. Therefore it is clear that both N_2 and N_3 lie in the same connected component of $\mathfrak{g}_0'$. Hence $k=2$ in this case. This shows that

$$\langle\eta^r,\eta^r\rangle T'(f) = \sum_{1\leqq i\leqq 3} c_i' \int_{N_i} \eta^r\partial(\eta^r) f\, dX \qquad (f \in C_c^\infty(N_0))$$

(in the notation of Theorem 1) where c_1', c_2', c_3' are certain constants and $c_2'=c_3'$ in case $[\Xi: LB]=2$. Furthermore, $\partial(\xi)T'=0$ as we have already observed above. Hence it follows from Theorem 1 that $c_1'=c_2'=c_3'$, and therefore

$$\langle\eta^r,\eta^r\rangle T'(f) = c_1' \int_N \eta^r\partial(\eta^r) f\, dX = c_1'\langle\eta^r,\eta^r\rangle \int_{\mathfrak{g}_0} f\, dX$$

for all $f \in C_c^\infty(N_0)$. This proves that T' coincides with c_1' on N_0.

We can now prove the following theorem.

THEOREM 2. *Let $\mathfrak{h}_0$ be a Cartan subalgebra of $\mathfrak{g}_0$ such that*[3] $\theta(\mathfrak{h}_0)=\mathfrak{h}_0$ *and $\mathfrak{h}_1$ a connected component of $\mathfrak{h}_0'=\mathfrak{h}_0 \cap \mathfrak{g}_0'$. For any $f \in \mathscr{C}(\mathfrak{g}_0)$, put*

$$\phi_f(H) = \pi(H) \int_{G^*} f(x^*H)\, dx^* \qquad (H \in \mathfrak{h}_0')$$

in the notation of Theorem 3 of [4(e)]. Then there exists a real number c such that

$$\lim_{H\to 0} \phi_f(H;\partial(\pi)) = cf(0) \qquad (H \in \mathfrak{h}_1)$$

for all $f \in \mathscr{C}(\mathfrak{g}_0)$. Moreover $c=0$ if $\mathfrak{h}_0$ is not fundamental.

[3] In view of the Corollary to Lemma 1 of [4(c)], the assumption $\theta(\mathfrak{h}_0)=\mathfrak{h}_0$ is obviously unnecessary. However, we make it here for convenience.

Since $\mathfrak{g}_0' \subset S$, it follows from Lemma 14 that $T' = c'$ on $\mathfrak{g}_0'$. Hence we conclude from the Corollary to Lemma 29 of [4(e)] that $T' = c'$ on $\mathfrak{g}_0$. In view of the definition of T', this implies that

$$\operatorname*{Lim}_{\mathfrak{h}_1} \phi_f(0;\partial(\pi)) = cf(0) \qquad (f \in \mathscr{C}(\mathfrak{g}_0)),$$

where c is a complex number independent of f. We know that η takes only real values on $\mathfrak{g}_0$ (see [4(e), § 6]). Select an element $H_0 \in \mathfrak{h}_0'$ and put $\epsilon = |\pi(H_0)|/\pi(H_0)$. Then it is clear that $\epsilon\pi$ takes only real values on $\mathfrak{h}_0'$. Let us choose a real-valued function $f \in \mathscr{C}(\mathfrak{g}_0)$ such that $f(0) = 1$. Then obviously $\partial(\epsilon\pi)(\epsilon\phi_f) = \epsilon^2\partial(\pi)\phi_f$ takes only real values on $\mathfrak{h}_0'$. Since $\epsilon^2 = \pm 1$, it follows that $c = \operatorname*{Lim}_{\mathfrak{h}_1} \phi_f(0;\partial(\pi))$ is real. Finally we know from the Corollary to Lemma 34 of [4(e)] that $c = 0$ if $\mathfrak{h}_0$ is not fundamental.

5. Proof of Theorem 3. Our next object is to prove the following theorem.

THEOREM 3. *Let $\mathfrak{h}_0$ be a fundamental Cartan subalgebra of $\mathfrak{g}_0$ such that*[3] *$\theta(\mathfrak{h}_0) = \mathfrak{h}_0$. Let $\mathfrak{h}_{(1)}, \cdots, \mathfrak{h}_{(q)}$ be all the distinct connected components of $\mathfrak{h}_0' = \mathfrak{h}_0 \cap \mathfrak{g}_0'$ and $c_1, \cdots, c_q$ the corresponding real numbers of Theorem 2. Then $c_1 + c_2 + \cdots + c_q \neq 0$.*

If the ranks of $\mathfrak{g}$ and $\mathfrak{k}$ are equal, this follows immediately from Lemma 41 of [4(e)] and Theorem 2. So let us now suppose that $\operatorname{rank} \mathfrak{g} > \operatorname{rank} \mathfrak{k}$. We use the notation of [4(e), § 5] and put $\mathfrak{g}_1 = \bigcup_{x \in G} x\mathfrak{h}_0'$. Then if the invariant measure dx^* on G/A is suitably normalized,

$$\int_{\mathfrak{g}_1} f(X)\,dX = \int_{G^* \times \mathfrak{h}_0'} f(x^*H)\,|\pi(H)|^2\,dx^*dH$$

for $f \in C_c(\mathfrak{g}_1)$. Here dX and dH are the regular Euclidean measures on $\mathfrak{g}_0$ and $\mathfrak{h}_0$, respectively. On the other hand, it follows from the Corollary to Lemma 15 of [4(e)] that

$$\int_{G^*} f(x^*H)\,dx^* = |\pi_+(H)|^{-1} \int f(k(mH + Z))\,dk\,dm\,dZ$$

for $H \in \mathfrak{h}_0'$ and $f \in \mathscr{C}(\mathfrak{g}_0)$. Here $\pi_+ = \prod_{\alpha \in P_+} \alpha$, dk, dm are the Haar measures on K and M, respectively, and dZ is the Euclidean measure on $\mathfrak{n}_0$. Moreover $\int_K dk = 1$. Since $\mathfrak{h}_0$ is fundamental, no root in P can vanish identically on $\mathfrak{h}_0 \cap \mathfrak{k}_0$ (see Lemma 33 of [4(e)]). Therefore $\alpha \neq -\theta\alpha$ for $\alpha \in P_+$. On the other hand, since α and $-\theta\alpha$ coincide on $\mathfrak{h}_{\mathfrak{p}_0}$, it follows from the

definition of P and P_+ (see [4(e), § 5]) that $-\theta\alpha \in P_+$ whenever $\alpha \in P_+$. Moreover, since $\operatorname{conj} \alpha(H) = -\alpha(\theta(H))$ $(\alpha \in P, H \in \mathfrak{h}_0)$, it is clear that π_+ takes only nonnegative real values on $\mathfrak{h}_0$.

A being the Cartan subgroup of G corresponding to $\mathfrak{h}_0$, let A' denote the normalizer of A in G. Then if W is the Weyl group of $\mathfrak{g}$ with respect to $\mathfrak{h}$, $W' = A'/A$ can be regarded as a subgroup[4] of W. Thus W' is a finite group which operates on $G^* = G/A$ on the right and leaves the measure dx^* invariant.[4] Now select two functions α_0, β_0 in $C_c^\infty(G^*)$ and $C_c^\infty(\mathfrak{h}_0')$, respectively, such that

$$\int \alpha_0(x^*)\,dx^* = \int \beta_0(H)\,|\,\pi(H)\,|^2\,dH = 1.$$

Let w' denote the order of W' and put

$$\alpha(x^*) = (w')^{-1} \sum_{s \in W'} \alpha(x^*s), \qquad \beta = (w')^{-1} \sum_{s \in W'} \beta_0^{\,s},$$

where $\beta_0^{\,s}(H) = \beta_0(s^{-1}H)$ $(H \in \mathfrak{h}_0)$. Then it is clear (see [4(e), § 11]) that there exists a function $F \in C_c^\infty(\mathfrak{g}_1)$ such that $F(x^*H) = \alpha(x^*)\beta(H)$ $(x^* \in G^*, H \in \mathfrak{h}_0')$. Moreover,

$$\int_{\mathfrak{g}_0} F(X)\,dX = \int_{G^*\times\mathfrak{h}_0'} F(x^*H)\,|\,\pi(H)\,|^2\,dx^*dH$$
$$= \int \alpha(x^*)\,dx^* \int \beta(H)\,|\,\pi(H)\,|^2\,dH = 1.$$

Let $\tilde{F}$ denote the Fourier transform of F and put

$$\phi_{\tilde{F}}(H) = \pi(H) \int_{G^*} \tilde{F}(x^*H)\,dx^* \qquad (H \in \mathfrak{h}_0').$$

Then in view of what we have seen above,

$$\phi_{\tilde{F}}(H) = \pi'(H) \int \tilde{F}(k(mH + Z))\,dk\,dm\,dZ,$$

where $\pi' = \prod_{\alpha \in P'} \alpha$ and $P' = P_0 \cup P_-$. Define ζ, $\mathfrak{l}_0$ etc. as in Lemma 15 of [4(e)] and let $\mathfrak{h}_0''$ be the set of all points $H \in \mathfrak{h}_0$ where $\pi'(H) \neq 0$. Then it follows from the Corollary of Lemma 42 of [4(e)] that $\mathfrak{h}_0''$ and $\mathfrak{h}_0'$ have the same number of connected components. Let $\mathfrak{h}_{(i)}''$ denote the connected component of $\mathfrak{h}_0''$ containing $\mathfrak{h}_{(i)}$. Then $\mathfrak{h}_{(1)}'', \cdots, \mathfrak{h}_{(q)}''$ are all the distinct connected components of $\mathfrak{h}_0''$. For any $g \in \mathscr{B}(\mathfrak{l}_0)$, put

$$\psi_g(H) = \pi'(H) \int_M g(mH)\,dm \qquad (H \in \mathfrak{h}_0'').$$

---

[4] See Sections 6, 8 and 11 of [4(e)] for the proof of these facts.

Then it follows from Lemma 19 of [4(e)] that $\psi_g \in \mathcal{C}(\mathfrak{h}_0'')$ and $\partial(\pi_+)\psi_g = \psi_{\partial(\zeta)g}$. On the other hand, we know from Theorem 2 (applied to $\mathfrak{l}_0$) that there exist constants c_i' $(1 \leqq i \leqq q)$ such that

$$\lim_{H\to 0} \psi_g(H;\partial(\pi')) = c_i' g(0) \qquad (H \in \mathfrak{h}_{(i)}'')$$

for every $g \in \mathcal{C}(\mathfrak{l}_0)$. Moreover, since $\mathfrak{h}_0 \cap \mathfrak{k}_0$ is maximal abelian in $\mathfrak{m}_0$, we conclude from Lemma 41 of [4(e)] (applied to $[\mathfrak{l}_0, \mathfrak{l}_0]$) that $c_1' + \cdots + c_q' \neq 0$. Now put

$$\bar{F}_1(X) = \int_K \bar{F}(kX)\,dk \quad (X \in \mathfrak{g}_0) \quad \text{and} \quad g_1(L) = \int_{\mathfrak{n}_0} \bar{F}_1(L+Z)\,dZ \quad (L \in \mathfrak{l}_0).$$

Then $\phi_{\bar{F}} = \psi_{g_1}$, and therefore $\partial(\pi)\phi_{\bar{F}} = \partial(\pi')\psi_{\partial(\zeta)g_1}$ on $\mathfrak{h}_0'$. Hence

$$\lim_{H\to 0} \phi_{\bar{F}}(H;\partial(\pi)) = \lim_{H\to 0} \psi_{\partial(\zeta)g_1}(H;\partial(\pi')) = c_i' g_1(0;\partial(\zeta)) \qquad (H \in \mathfrak{h}_{(i)}).$$

Now put $F_1(X) = \int_K F(kX)\,dk$ $(X \in \mathfrak{g}_0)$. Then since $\mathfrak{g}_0$ is the orthogonal sum of $\mathfrak{n}_0$, $\mathfrak{l}_0$ and $\theta(\mathfrak{n}_0)$, we conclude from the theory of Fourier transforms that

$$\begin{aligned} g_1(L') &= \int_{\mathfrak{n}_0} \bar{F}_1(L'+Z)\,dZ \\ &= \int_{\mathfrak{l}_0\times\mathfrak{n}_0} \exp\{(-1)^{\frac{1}{2}} B(L',L)\} F_1(L+Z)\,dL\,dZ \qquad (L' \in \mathfrak{l}_0). \end{aligned}$$

Here dL is the (suitably normalized) Euclidean measure on $\mathfrak{l}_0$. Hence it follows easily that

$$g_1(0;\partial(\zeta)) = (-1)^{r_+/2} \int_{\mathfrak{l}_0\times\mathfrak{n}_0} \zeta(L) F_1(L+Z)\,dL\,dZ,$$

where r_+ is the number of roots in P_+. Select $L \in \mathfrak{l}_0$ and $Z \in \mathfrak{n}_0$ such that $\zeta(L)F_1(L+Z) \neq 0$. Then it follows from Lemma 16 of [4(e)] that $nL = L + Z = xH$ for some $n \in N$, $x \in G$ and $H \in \mathfrak{h}_0'$. Hence $L = yH$ if $y = n^{-1}x$. Let $\mathfrak{h}_L$ denote the centralizer of L in $\mathfrak{g}_0$. Then $\mathfrak{h}_L = y\mathfrak{h}_0$, and therefore $\mathfrak{h}_L$ is a fundamental Cartan subalgebra of $\mathfrak{g}_0$. On the other hand, since the ranks of $\mathfrak{l}_0$ and $\mathfrak{g}_0$ are the same, it is evident that $\mathfrak{h}_L \subset \mathfrak{l}_0$, and therefore, obviously, both $\mathfrak{h}_0$ and $\mathfrak{h}_L$ are fundamental Cartan subalgebras[5] of the reductive algebra $\mathfrak{l}_0$. Hence it follows from the Corollary to Lemma 32 of[4(e)] that $\mathfrak{h}_L = m\mathfrak{h}_0$ for some $m \in M$. Since L is a regular element

[5] The notion of a fundamental Cartan subalgebra is extended to reductive Lie algebras over R in the obvious way.

of $\mathfrak{h}_L$, we conclude that $L = mH_1$ for some $H_1 \in \mathfrak{h}_0'$, and therefore $\zeta(L) = \zeta(H_1) = \pi_+(H_1) > 0$. This shows that

$$g_1(0;\partial(\zeta)) = \epsilon \int |\zeta(L)| F_1(L+Z)\, dL dZ = \epsilon \int_{\mathfrak{l}_0 \times N} |\zeta(L)|^2 F_1(nL)\, dL dn$$

from Lemma 15 of [4(e)]. (Here $\epsilon = (-1)^{r_+/2}$.) Now put $\mathfrak{l}_1 = \bigcup_{m \in M} m(\mathfrak{h}_0')$. Then as we have seen above, $\zeta(L) F_1(nL) = 0$ $(n \in N, L \in \mathfrak{l}_0)$ unless $L \in \mathfrak{l}_1$. Moreover, it is clear that there exists a positive constant a such that

$$\int_{\mathfrak{l}_1} f(L)\, dL = a \int_{M \times \mathfrak{h}_0'} f(mH) |\pi'(H)|^2\, dm dH$$

for all $f \in C_c^\infty(\mathfrak{l}_1)$. Hence

$$\begin{aligned} g_1(0;\partial(\zeta)) &= a\epsilon \int |\zeta(H)|^2 F_1(nmH) |\pi'(H)|^2\, dm dH dn \\ &= a\epsilon \int |\pi(H)|^2 F(knmH)\, dk dm dH dn \\ &= a\epsilon \int |\pi(H)|^2 F(x^*H)\, dx^* dH \\ &= a\epsilon \int F(X)\, dX = a\epsilon \end{aligned}$$

from the Corollary to Lemma 14 of [4(e)]. This proves that

$$\lim_{H \to 0} \phi_{\tilde{F}}(H;\partial(\pi)) = a\epsilon c_i' \qquad (H \in \mathfrak{h}_{(i)}).$$

But since $\tilde{F}(0) = 1$, this implies that $c_i = a\epsilon c_i'$, and therefore $c_1 + \cdots + c_q = a\epsilon(c_1' + \cdots + c_q') \neq 0$.

6. Appendix. Let $\mathfrak{z}_0$ be a subalgebra of $\mathfrak{g}_0$ and Ξ_0 the analytic subgroup of G corresponding to $\mathfrak{z}_0$. We assume that $\mathfrak{z}_0$ is reductive in $\mathfrak{g}_0$. Let Ξ' be the normalizer of $\mathfrak{z}_0$ in G.

LEMMA 15. *Suppose* $\operatorname{rank} \mathfrak{z}_0 = \operatorname{rank} \mathfrak{g}_0$. *Then the factor group* Ξ'/Ξ_0 *is finite.*

Since $\mathfrak{z}_0$ is reductive, we can select (see the Corollary to Lemma 2 of [4(c)]) a finite number of Cartan subalgebras $\mathfrak{h}_1, \cdots, \mathfrak{h}_r$ of $\mathfrak{z}_0$ such that every Cartan subalgebra of $\mathfrak{z}_0$ is of the form $x\mathfrak{h}_i$ for some $x \in \Xi_0$ and some i $(1 \leqq i \leqq r)$. Suppose $\mathfrak{h}_1$ is conjugate to $\mathfrak{h}_i$ $(1 \leqq i \leqq k)$ under Ξ' but not to $\mathfrak{h}_i$ $(k < i \leqq r)$. Select $\xi_i \in \Xi'$ such that $\xi_i \mathfrak{h}_1 = \mathfrak{h}_i$ $1 \leqq i \leqq k$ $(\xi_1 = 1)$. Now let $x \in \Xi'$. Then it is clear that $x\mathfrak{h}_1 = y\mathfrak{h}_i = y\xi_i\mathfrak{h}_1$ for some $y \in \Xi_0$ and

some i $(1 \leqq i \leqq k)$. Since rank $\mathfrak{z}_0 = $ rank $\mathfrak{g}_0$ and $\mathfrak{z}_0$ is reductive in $\mathfrak{g}_0$, $\mathfrak{h}_1$ is also a Cartan subalgebra of $\mathfrak{g}_0$. Let A' be the normalizer of $\mathfrak{h}_1$ in G and A the Cartan subgroup of G corresponding to $\mathfrak{h}_1$. Then A'/A is a finite group (see [4(c), Lemma 10]). On the other hand, if A_0 is the analytic subgroup of G corresponding to $\mathfrak{h}_1$, $A \cap \Xi_0 \supset A_0$. Moreover, from Lemma 7 of [4(e)], A/A_0 is both compact and discrete and hence finite. This proves that $(A' \cap \Xi')/A_0$ is a finite group. Select elements $a_j \in A' \cap \Xi'$ such that $A' \cap \Xi'$ is the union of $a_j A_0$ $(1 \leqq j \leqq N)$. Then it is clear that $x^{-1} y \xi_i \in a_j A_0$ for some j, and therefore $x \in y \xi_i A_0 a_j^{-1} \subset \xi_i a_j^{-1} \Xi_0$ since Ξ_0 is obviously normal in Ξ'. This shows that the order of the group Ξ'/Ξ_0 cannot exceed rN.

Let $\mathfrak{l}$ be the semisimple Lie algebra of dimension 3 spanned over by the three elements H, X, Y which satisfy the relations $[H,X] = 2X$, $[H,Y] = -2Y$, $[X,Y] = H$. Let Ω denote the Casimir polynomial of $\mathfrak{l}$. Then

$$\Omega(tH + xX + yY) = 8(t^2 + xy) \qquad (t, x, y \in C).$$

Moreover, $\mathfrak{h} = CH$ is a Cartan subalgebra of $\mathfrak{l}$ and there are only two roots α and $-\alpha$ given by $\alpha(H) = 2H$. We take α as the positive root. Then Ω coincides with $2\alpha^2$ on $\mathfrak{h}$. Hence an element $Z \in \mathfrak{l}$ is regular if and only if $\Omega(Z) \neq 0$. Let $\mathfrak{l}'$ denote the set of regular elements in $\mathfrak{l}$.

Consider the automorphism ν of $\mathfrak{l}$ given by[2]

$$\nu = \exp\{-(-1)^{\frac{1}{2}}(\pi/4)\operatorname{ad}(X+Y)\}.$$

A simple calculation shows that $\nu(H) = (-1)^{\frac{1}{2}}(X-Y)$, $\nu(X-Y) = (-1)^{\frac{1}{2}}H$ and $\nu(X+Y) = X+Y$. Therefore $RH + RX + RY$ and $\nu(RH + RX + RY) = R(-1)^{\frac{1}{2}}H + R(-1)^{\frac{1}{2}}(X-Y) + R(X+Y)$ are two isomorphic real forms of $\mathfrak{l}$. Since $\Omega(H) = 8 > 0$, they are not compact. On the other hand $R(-1)^{\frac{1}{2}}H + R(X-Y) + R(-1)^{\frac{1}{2}}(X+Y)$ is also a real form of $\mathfrak{l}$ which is compact since Ω is negative definite on it. We shall denote by $\mathfrak{l}_0$ either one of the two[6] real forms $RH + RX + RY$ and $R(-1)^{\frac{1}{2}}H + R(X-Y) + R(-1)^{\frac{1}{2}}(X+Y)$. Put $\mathfrak{l}_0' = \mathfrak{l}' \cap \mathfrak{l}_0$. If $\mathfrak{l}_0$ is compact, zero is the only singular point in $\mathfrak{l}_0$, and therefore $\mathfrak{l}_0'$ is connected. On the other hand, if $\mathfrak{l}_0$ is noncompact, the singular set of $\mathfrak{l}_0$ consists of all points $tH + xX + yY$ $(t, x, y \in R)$ satisfying the equation $t^2 + xy = 0$. In this case $\mathfrak{l}_0'$ has three connected components $\mathfrak{l}_1, \mathfrak{l}_2, \mathfrak{l}_3$ given as follows:

$$\mathfrak{l}_1 : t^2 + xy > 0; \quad \mathfrak{l}_2 : t^2 + xy < 0, x > y; \quad \mathfrak{l}_3 : t^2 + xy < 0, x < y.$$

Notice that $H \in \mathfrak{l}_1$ and $X - Y \in \mathfrak{l}_2$. Let U be a neighborhood of zero in $\mathfrak{l}_0$.

---

[6] It is not difficult to see that apart from isomorphism these are the only two real forms of $\mathfrak{l}$. However, we do not need this fact.

We assume that $tZ \in U$ $(0 \leqq t \leqq 1)$ whenever $Z \in U$. Then it is obvious that $U \cap \mathfrak{l}_0'$ is connected if $\mathfrak{l}_0$ is compact and consists of three connected components $U \cap \mathfrak{l}_i$ $(i = 1, 2, 3)$ if $\mathfrak{l}_0$ is not compact.

From now on we consider only the noncompact case. The mapping $(tH + xX + yY) \to \begin{pmatrix} t & x \\ y & -t \end{pmatrix}$ $(t, x, y \in C)$ is an isomorphism of $\mathfrak{l}$ onto the Lie algebra of all 2×2 complex matrices with trace zero and we may identify the two under this isomorphism. Then $\mathfrak{l}_0$ becomes the real subalgebra of $\mathfrak{l}$ consisting of the real matrices. It is clear that $\Omega(Z) = -8 \det Z$ and the two eigenvalues of Z are $\pm(-\det Z)^{\frac{1}{2}}$ $(Z \in \mathfrak{l})$. Call an element $Z \in \mathfrak{l}_0'$ hyperbolic or elliptic according as these eigenvalues are real or pure imaginary. Then $\mathfrak{l}_1$ consists of hyperbolic and $\mathfrak{l}_2 \cup \mathfrak{l}_3$ of elliptic elements. Let L be the (connected) adjoint group of $\mathfrak{l}_0$. It is easily seen that every hyperbolic element is conjugate under L to tH for some real $t \neq 0$. Moreover, since $\exp \frac{\pi}{2} \operatorname{ad}(X - Y)$ maps H into $-H$, we can always assume that $t > 0$. Let R^+ and R^- be the set of all positive and negative real numbers, respectively, and put $R' = R^+ \cup R^-$. Then we prove similarly that every element in $\mathfrak{l}_2$ and $\mathfrak{l}_3$, respectively, is conjugate to $\theta(X - Y)$ for some θ in R^+ and R^-. It is easy to check that there exists no element $x \in L$ such that $x(X - Y) = -(X - Y)$.

Let dZ denote the Euclidean measure on $\mathfrak{l}_0$ given by[2] $dZ = (2\pi)^{-1} dt\,dx\,dy$, where $Z = tH + xX + yY$ $(t, x, y \in R)$. Consider the one parameter subgroup K in L corresponding to $(X - Y)$. Then K is compact. Let dk denote the normalized Haar measure on K so that $\int_K dk = 1$.

LEMMA 16. *For any $f \in C_c(\mathfrak{l}_0)$ put*

$$\Phi_f(t) = \int_{-\infty}^{\infty} \bar{f}(tH + xX)\,dx \qquad (t \in R^+)$$

$$\Psi_f(\theta) = |\theta| \int_0^{\infty} \bar{f}(\theta(e^t X - e^{-t} Y))(e^t - e^{-t})\,dt \qquad (\theta \in R'),$$

where $\bar{f}(Z) = \int f(kZ)\,dk$ $(Z \in \mathfrak{l}_0)$. Then

$$\int_{\mathfrak{l}_1} f(Z)\,dZ = \int_0^{\infty} t\Phi_f(t)\,dt$$

$$\int_{\mathfrak{l}_2} f(Z)\,dZ = \int_0^{\infty} \theta\Psi_f(\theta)\,d\theta, \qquad \int_{\mathfrak{l}_3} f(Z)\,dZ = \int_{-\infty}^{0} |\theta|\,\Psi_f(\theta)\,d\theta.$$

Since f vanishes outside a compact set, it is seen without difficulty

that Φ_f and Ψ_f are well defined. Put $k_\theta = \exp\{\theta \operatorname{ad}(X-Y)\}$ $(\theta \in R)$ and consider the mapping $(\theta, t, x) \to (t', x', y')$ given by

$$k_\theta(tH + xX) = t'H + x'X + y'Y,$$

where $0 \leqq \theta < \pi$, $t > 0$ and $x \in R$. It follows from general considerations (see [4(a), Lemma 9]) that

$$|\partial(t', x', y')/\partial(\theta, t, x)| = ct,$$

where c is a constant. To obtain its value we compute the Jacobian at $\theta = 0$. Now

$$\{(\partial/\partial\theta) k_\theta(tH + xX)\}_{\theta=0} = [X - Y, tH + xX] = xH - 2tX - 2tY.$$

Hence

$$\left\{\frac{\partial(t', x', y')}{\partial(\theta, t, x)}\right\}_{\theta=0} = \det\begin{pmatrix} x & -2t & -2t \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = -2t.$$

Therefore $c = 2$. Let A, N be the one-parameter subgroups of L corresponding to H and X, respectively. Then we know from general considerations (see Iwasawa [5]) that $L = KNA$. Therefore every element in $\mathfrak{l}_1$ can be written in the form $t(knH)$ $(k \in K, n \in N, t \in R^+)$. Since $k_\pi = 1$, it follows that every point of $\mathfrak{l}_1$ is obtained under our mapping $(\theta, t, x) \to k_\theta(tH + xX)$. Moreover, a direct computation shows that this mapping is univalent. Hence

$$\int_{\mathfrak{l}_1} f(Z)\, dZ = 2(2\pi)^{-1} \int_0^\pi d\theta \int_0^\infty t\,dt \int_{-\infty}^\infty f(k_\theta(tH + xX))\, dx$$
$$= \int_0^\infty t\,dt \int_{-\infty}^\infty \bar{f}(tH + xX)\, dx = \int_0^\infty t\Phi_f(t)\, dt.$$

Moreover, if $L^* = L/A$ and dx^* is the invariant measure on L^*, it follows from the Corollary to Lemma 15 of [4(e)] that

$$\Phi_f(t) = t \int_{L^*} f(tx^*H)\, dx^* \qquad (f \in C_c(\mathfrak{l}_0), t \in R^+)$$

provided dx^* is suitably normalized.

Now put $a_t = \exp\{t \operatorname{ad} H\}$ $(t \in R)$ and consider the mapping $(\phi, t, \theta) \to (t', x', y')$ given by

$$k_\phi a_t(\theta(X - Y)) = t'H + x'X + y'Y$$

$(0 \leqq \phi < \pi, t \in R^+, \theta \in R')$. Then again it follows by general considerations[7] (or by actual computation) that

$$|\partial(t', x', y')/\partial(\phi, t, \theta)| = c\theta^2(e^{2t} - e^{-2t}),$$

where c is a constant. In order to obtain c we evaluate the left side at $\phi = 0$. Then $\theta a_t(X - Y) = \theta(e^{2t}X - e^{-2t}Y)$. Hence

$$[(\partial/\partial\phi)\{k_\phi a_t(\theta(X - Y))\}]_{\phi=0} = \theta[X - Y, e^{2t}X - e^{-2t}Y] = \theta(e^{2t} - e^{-2t})H.$$

Therefore

$$\left\{\frac{\partial(t', x', y')}{\partial(\phi, t, \theta)}\right\}_{\phi=0} = \det\begin{pmatrix} \theta(e^{2t} - e^{-2t}) & 0 & 0 \\ 0 & 2\theta e^{2t} & 2\theta e^{-2t} \\ 0 & e^{2t} & -e^{-2t} \end{pmatrix}$$

$= -4\theta^2(e^{2t} - e^{-2t})$, and so $c = 4$. A_+ be the set of all a_t with $t \geqq 0$. It is easy to see that $L = KA_+K$. Therefore every element in $\mathfrak{l}_2(\mathfrak{l}_3)$ can be written in the form $k_\phi a_t(\theta(X - Y))$ with $0 \leqq \phi < \pi$, $t \geqq 0$, $\theta \in R^+$ ($\theta \in R^-$ respectively). Moreover, if $t \neq 0$, this can be done only in one way since the normalizer of $R(X - Y)$ in L is exactly K. Hence

$$\int_{\mathfrak{l}_2} f(Z)\,dZ = 4(2\pi)^{-1}\int_0^\pi d\phi \int_0^\infty \theta^2 d\theta \int_0^\infty f(k_\phi a_t(\theta(X - Y)))(e^{2t} - e^{-2t})\,dt$$
$$= \int_0^\infty \theta\Psi_f(\theta)\,d\theta.$$

Similarly,

$$\int_{\mathfrak{l}_3} f(Z)\,dZ = 4(2\pi)^{-1}\int_0^\pi d\phi \int_0^\infty \theta^2 d\theta \int_0^\infty f(k_\phi a_t(-\theta(X - Y)))(e^{2t} - e^{-2t})\,dt$$
$$= \int_{-\infty}^0 |\theta|\,\Psi_f(\theta)\,d\theta.$$

On the other hand, if dz is the Haar measure on L, it follows from [4(b), Lemma 38] that

$$\Psi_f(\theta) = |\theta|\int_L f(\theta z(X - Y))\,dz \qquad (f \in C_c(\mathfrak{l}_0))$$

if dz is suitably normalized.

Now suppose $f \in C_c^\infty(\mathfrak{l}_0)$. Then we can conclude from Theorem 3 of [4(e)] that $\Phi_f \in \mathcal{C}(R^+)$ and $\Psi_f \in \mathcal{C}(R')$.

---

[7] See [4(a), p. 501] and Lemma 38 of [4(b)].

LEMMA 17. *Let f be a function in $C_c^\infty(\mathfrak{l}_0)$. Then*

$$\Phi_{\partial(\Omega)f}(t) = \tfrac{1}{8}\, d^2\Phi_f(t)/dt^2 \qquad (t \in R^+)$$

$$\Psi_{\partial(\Omega)f}(t) = -\tfrac{1}{8}\, d^2\Psi_f(\theta)/d\theta^2 \qquad (\theta \in R').$$

Since $\Omega(H) = 8$ and $\Omega(X - Y) = -8$, this is an immediate consequence of Theorem 3 of [4(e)].

Put $\Phi_f(0) = \lim_{t\to 0} \Phi_f(t)$ $(t \in R^+)$, $\Psi_f^+(0) = \lim_{\theta\to 0} \Psi_f(\theta)$ $(\theta \in R^+)$ and $\Psi_f^-(0) = \lim_{\theta\to 0} \Psi_f(\theta)$ $(\theta \in R^-)$ for $f \in C_c^\infty(\mathfrak{l}_0)$. (We know from Lemma 25 of [4(e)] that all these limits exist). Then it is obvious that $\Phi_f(0) = \int_{-\infty}^{\infty} \bar{f}(xX)\, dx$. So now let us consider $\Psi_f^\pm(0)$. It is obvious that

$$\lim_{\theta\to 0} \theta \int_0^\infty \bar{f}(\theta(e^tX - e^{-t}Y))e^{-t}\, dt = 0.$$

On the other hand, if $\theta \in R^+$,

$$\theta \int_0^\infty \bar{f}(\theta(e^tX - e^{-t}Y))e^t\, dt = \int_\theta^\infty \bar{f}(tX - \theta^2 t^{-1}Y)\, dt.$$

Moreover, if M is an upper bound for $|\partial(Y)\bar{f}|$, it is clear that $|\bar{f}(tX + sY) - \bar{f}(tX)| \leqq |s|\, M$ $(s \in R)$, and since $\bar{f}$ vanishes outside a compact set, we can select $T \in R^+$ such that $\bar{f}(tX + sY) = 0$ for all $s \in R$ if $t \geqq T$. Then if $\theta < T$, it follows that

$$\int_\theta^\infty |\bar{f}(tX - \theta^2 t^{-1}Y) - \bar{f}(tX)|\, dt \leqq \theta^2 \int_\theta^T Mt^{-1}\, dt = M\theta^2 \log(T/\theta).$$

Hence

$$\lim_{\theta\to 0} \int_0^\infty \bar{f}(tX - \theta^2 t^{-1}Y)\, dt = \int_0^\infty \bar{f}(tX)\, dt \qquad (\theta \in R^+).$$

This proves that $\Psi_f^+(0) = \int_0^\infty \bar{f}(tX)\, dt$. Similarly one shows that $\Psi_f^-(0) = \int_{-\infty}^0 \bar{f}(tX)\, dt$. Thus we have obtained the following result.

LEMMA 18. *For any $f \in C_c^\infty(\mathfrak{l}_0)$,*

$$\Phi_f(0) = \int_{-\infty}^\infty \bar{f}(tX)\, dt, \quad \Psi_f^+(0) = \int_0^\infty \bar{f}(tX)\, dt, \quad \Psi_f^-(0) = \int_{-\infty}^0 \bar{f}(tX)\, dt.$$

COROLLARY. *Put $f_m = \partial(\Omega^m)(\Omega^m f)$ $(f \in C_c^\infty(\mathfrak{l}_0), m \geqq 0)$. Then*

$$\Phi_{f_m}(0) = c_m \Phi_f(0), \quad \Psi_{f_m}^+(0) = c_m \Psi_f^+(0), \quad \Psi_{f_m}^-(0) = c_m \Psi_f^-(0),$$

where $c_m = 4^m m!\, \Gamma(m + \tfrac{1}{2})/\Gamma(\tfrac{1}{2})$.

It is obvious that the differential operator $\partial(\Omega^m)\circ\Omega^m$ is invariant under L and therefore also under K. Moreover $\Omega(X)=0$. Hence our assertion is an immediate consequence of the Corollary to Lemma 19 of [4(d)].

Let U be an open neighborhood of zero in $\mathfrak{l}_0$.

LEMMA 19. *Suppose c_1, c_2, c_3 are three constants and*

$$c_1\Phi_f(0)=c_2\Psi_f^+(0)+c_3\Psi_f^-(0)$$

for all $f\in C_c^\infty(U)$. Then $c_1=c_2=c_3$.

Let $\mathfrak{l}^+$ be the set of all points of the form $tH+xX+yY$ $(t,y\in R;x\in R^+)$. We can obviously choose a positive number x_0 and a real-valued nonnegative function $f\in C_c^\infty(U\cap\mathfrak{l}^+)$ such that $f(x_0X)=1$. A simple calculation shows that

$$k_\theta X=(\cos\theta\sin\theta)H+(\cos\theta)^2X-(\sin\theta)^2Y \qquad (\theta\in R).$$

Hence it is clear from Lemma 18 that $\Phi_f(0)=\Psi_f^+(0)>0$ while $\Psi_f^-(0)=0$. Therefore it follows from our hypothesis that $c_1=c_2$. Similarly we prove that $c_1=c_3$.

Let E be a vector space over R of dimension n and let $d\mu$ denote the Euclidean measure on E.

LEMMA 20. *Let U be a nonempty open set in E and $p_1,\cdots,p_r$ a finite number of continuous functions on U. Suppose $p_1,\cdots,p_r$ are linearly independent over C. Then we can select a function $f\in C_c^\infty(U)$ such that*

$$\int_U fp_i\,d\mu=\begin{cases}1 \text{ if } i=1,\\ 0 \text{ if } i\neq 1\end{cases} \qquad (1\leqq i\leqq r).$$

Let τ_i denote the distribution on U given by

$$\tau_i(g)=\int_U gp_i\,d\mu \qquad (g\in C_c^\infty(U)).$$

Let V be a vector space over C of dimension r and let $v_1,\cdots,v_r$ be a base for V. For any $g\in C_c^\infty(U)$, put $\tau(g)=\sum_{1\leqq i\leqq r}\tau_i(g)v_i$ and let V_1 be the subspace of V consisting of all elements of the form $\tau(g)$. We claim $V_1=V$. For otherwise, we could select a linear function $\lambda\neq 0$ on V such that $\lambda(\tau(g))=0$ for all $g\in C_c^\infty(U)$. Put $\lambda_i=\lambda(v_i)$ $1\leqq i\leqq r$. Then not every λ_i is zero and

$$\lambda(\tau(g))=\sum_{1\leqq i\leqq r}\lambda_i\tau_i(g)=0 \qquad (g\in C_c^\infty(U)).$$

Let $p = \sum_{1 \leqq i \leqq r} \lambda_i p_i$. Since p_i are linearly independent, $p \neq 0$, and therefore it is obvious that $\int pg\, d\mu \neq 0$ for some $g \in C_c^\infty(U)$. However

$$\int pg\, d\mu = \sum_{1 \leqq i \leqq r} \lambda_i \tau_i(g) = \lambda(\tau(g)) = 0,$$

and so we get a contradiction. This proves that $V_1 = V$, and therefore we can choose $f \in C_c^\infty(U)$ such that $\tau(f) = v_1$. Then obviously f fulfills the desired condition.

INSTITUTE FOR ADVANCED STUDY,
COLUMBIA UNIVERSITY.

REFERENCES.

[1] E. Cartan, *Ann. Ecole Norm. Sup.*, vol. 44 (1927), pp. 345-467.

[2] C. Chevalley, *Theory of Lie groups*, Princeton University Press, 1946.

[3] L. Garding, *Math. Scand.*, vol. 1 (1953), pp. 55-72.

[4] Harish-Chandra, (a) *Trans. Amer. Math. Soc.*, vol. 76 (1954), pp. 485-528.
(b) *Amer. Jour. Math.*, vol. 78 (1956), pp. 564-628.
(c) *Trans. Amer. Math. Soc.*, vol. 83 (1956), pp. 98-163.
(d) *Amer. Jour. Math.*, vol. 79 (1957), pp. 87-120.
(e) *Amer. Jour. Math.*, vol. 79 (1957), pp. 193-257.
(f) *Proc. Nat. Acad. Sci. U. S. A.*, vol. 42 (1956), pp. 538-540.

[5] K. Iwasawa, *Ann. of Math.*, vol. 50 (1949), pp. 507-557.

[6] F. John, *Proceedings of the symposium on spectral theory and differential problems*, Stillwater, Okla., (1951), pp. 113-175.

[7] L. Schwartz, *Théorie des distributions I*, Paris, Hermann, 1950.

[8] A. Weil, *L'intégration dans les groupes topologiques et ses applications*, Paris, Hermann, 1940.

LOCAL CHARACTERIZATION OF INTEGRAL QUADRATIC FORMS BY GAUSS SUMS.*

By O. T. O'Meara.

This paper is concerned with the existence of a solution X to the $n \times n$ matrix equation $X^T H X = G$, with H and G symmetric. When the coefficient field has characteristic $\neq 2$, this is the same as determining the equivalence class of a quadratic form. Invariants for the equivalence class of H under the transformation $X^T H X$ have been given over certain arithmetical fields and a discussion of these results can be found, for instance, in Witt's work on quadratic forms [11]. The problem of finding when there is a unimodular solution X has been solved over local fields by Durfee, Jones, Minkowski and the author. As far as the known results go, the general solution to this problem is in terms of Hasse symbols. There is another characterization due to Minkowski which involves Gauss sums but, to use the language of valuation theory, this has been proved over the p-adic numbers only. We shall extend this result to local fields with finite residue class field.

There are three levels of difficulty in the local theory and they are determined by the way in which 2 splits in the coefficient field. Thus we shall find that the Gauss sums alone characterize the integral equivalence class of H when $\operatorname{ord} 2 = 0$, while one must include the type as an invariant in the unramified case. This much is in agreement with the rational p-adic results of Minkowski. In the ramified theory further conditions must be imposed and these will be given in terms of the invariants a_j, $v(j)$ introduced in [8]; however, we shall see that the Gauss sums alone form a complete set of invariants for a unitary H over any of these fields.

1. Group characters. We consider a local field F with ring of integers $\mathfrak{o}$ and prime ideal $\pi\mathfrak{o}$. We assume that the residue class field $\mathfrak{o}/(\pi\mathfrak{o})$ is finite of characteristic 2, while F itself has characteristic 0. Thus the number of elements in $\mathfrak{o}/(\pi\mathfrak{o})$ is of the form 2^f. Define the ramification index e by the relation $e = \operatorname{ord} 2$; under these restrictions we see that $1 \leqq e < \infty$. It follows [1] from these assumptions that F must be a finite extension of the 2-adic numbers, the degree of the extension being ef. The $\mathfrak{o}$-ideal $\mathfrak{o}(r)$ is

* Received December 4, 1956.

defined by $\mathfrak{o}(r)=\pi^r\mathfrak{o}$ for $r\geqq 0$. If $q\geqq r$, then $\mathfrak{o}(q)\subseteq\mathfrak{o}(r)$. If $\alpha_1,\cdots$ is a complete set of representatives of $\mathfrak{o}/\mathfrak{o}(r)$ and if $\beta_1,\cdots$ is a complete set for $\mathfrak{o}/\mathfrak{o}(q-r)$, then it is easily seen that $\alpha_i+\pi^r\beta_j$ form a complete set for $\mathfrak{o}/\mathfrak{o}(q)$. We shall write $[\pi^t]$ to denote an element of ordinal t; $\{\pi^t\}$ will denote an element of ordinal $\geqq t$. The relation $\alpha\cong\beta$ means that there is a unit in $\mathfrak{o}$ whose square in α/β. If ϵ is a unit, then $\epsilon\cong 1 \bmod \pi$ since $\mathfrak{o}/(\pi\mathfrak{o})$ is a finite field. Consider the greatest value of k for which $\epsilon\cong 1+[\pi^k]$. If $k>2e$, then $k=\infty$ by Hensel's lemma. On the other hand,

$$k<2e\Longrightarrow k \text{ odd}. \tag{1}$$

A mapping χ of $\mathfrak{o}$ into the multiplicative group of complex numbers is called a *character* if $\chi(\alpha+\beta)=\chi(\alpha)\chi(\beta)$ for all α,β in $\mathfrak{o}$, and if $\chi(\mathfrak{o}(r))=1$ for some $r<\infty$. Any ideal $\mathfrak{o}(r)$ for which $\chi(\mathfrak{o}(r))=1$ is called a *support* of χ; the largest ideal in $\mathfrak{o}$ which supports χ is called the *maximal support* of χ; $\mathfrak{o}(m(\chi))$, or simply $\mathfrak{o}(m)$, will be used to denote the maximal support of χ; then we have $0\leqq m(\chi)<\infty$. For example, $m(\chi)=0$ means that χ is the identity character taking the value 1 for all $\alpha\in\mathfrak{o}$. Note that $\chi(0)=1$ for any χ.

We define the *character group* $\mathfrak{D}$ of $\mathfrak{o}$ as the set of all characters on $\mathfrak{o}$, the group structure being defined by

$$\chi\theta(\alpha)=\chi(\alpha)\theta(\alpha) \qquad \text{for all } \alpha\in\mathfrak{o};$$

the identity of $\mathfrak{D}$ is the identity character. If $\mathfrak{D}(r)\subseteq\mathfrak{D}$ consists of all $\chi\in\mathfrak{D}$ for which $\chi(\mathfrak{o}(r))=1$, then $\mathfrak{D}(r)$ is a subgroup of $\mathfrak{D}$. And any character $\chi\in\mathfrak{D}(r)$ can be interpreted as a character on $\mathfrak{o}/\mathfrak{o}(r)$ since χ is supported by $\mathfrak{o}(r)$. Conversely, every character on $\mathfrak{o}/\mathfrak{o}(r)$ can be interpreted as a character in $\mathfrak{D}(r)$. Since $\mathfrak{o}/\mathfrak{o}(r)$ is a finite abelian group of order 2^{fr}, it follows from the theory of characters [1] that $\mathfrak{D}(r)$ contains 2^{fr} elements. We have the reverse order relations

$$\mathfrak{o}=\mathfrak{o}(0)\supset\mathfrak{o}(1)\cdots\supset\mathfrak{o}(r)\supset\mathfrak{o}(r+1)\cdots\supset 0,$$

$$1=\mathfrak{D}(0)\subset\mathfrak{D}(1)\cdots\subset\mathfrak{D}(r)\subset\mathfrak{D}(r+1)\cdots\subset\mathfrak{D}.$$

From the theory of characters on a finite abelian group [10] we get

$$\sum_{\alpha \bmod \mathfrak{o}(r)}\chi(\alpha)=\begin{cases}0 & \text{if } \chi\neq 1,\ \chi\in\mathfrak{D}(r)\\ 2^{fr} & \text{if } \chi=1.\end{cases} \tag{2}$$

Consider a character χ with maximal support $\mathfrak{o}(m)$. For fixed integers A and B define

$$\theta(\alpha)=\chi(A\alpha),\qquad \psi(\alpha)=\chi(B\alpha) \qquad \text{for all } \alpha\in\mathfrak{o}.$$

Then $\theta \in \mathfrak{D}(m)$, $\psi \in \mathfrak{D}(m)$. It is easy to see that

$$\theta = \psi \Leftrightarrow A \equiv B \bmod \mathfrak{o}(m).$$

Thus 2^{fm} characters can be obtained by allowing A to vary through $\mathfrak{o}$; in other words, we then get all $\mathfrak{D}(m)$. If $A = E\pi^k$ $(0 \leqq k \leqq m)$ and if E is allowed to vary through all units mod π^{m-k}, the dependent variable θ ranges through all those characters with maximal support $\mathfrak{o}(m-k)$. We define $\chi_{(k)}$ as the character given by

$$\chi_{(k)}(\alpha) = \chi(\pi^k\alpha) \qquad \text{for all } \alpha \in \mathfrak{o}. \tag{3}$$

Now let θ be any character with maximal support $\mathfrak{o}(m-k)$; then $\theta(\alpha) = \chi(\pi^k E\alpha)$ for some fixed unit E; so putting $\psi(\alpha) = \chi(E\alpha)$ we see that $\psi_{(k)}(\alpha) = \psi(\pi^k\alpha) = \theta(\alpha)$ for all $\alpha \in \mathfrak{o}$. So we have proved that *every θ with maximal support $\mathfrak{o}(m-k)$ is a $\psi_{(k)}$ for some ψ with maximal support $\mathfrak{o}(m)$.* We shall need the dual of (2):

$$\sum_{\chi \in \mathfrak{D}(r)} \chi(\alpha) = \begin{cases} 0 & \text{if } \alpha \notin \mathfrak{o}(r) \\ 2^{fr} & \text{if } \alpha \in \mathfrak{o}(r). \end{cases} \tag{4}$$

Remark. Our given local field must contain a replica of the 2-adic numbers Q_2; in fact assume that $Q_2 \subseteq F$. Any $A \in Q_2$ can be put in the form $A = \lambda 2^{-\mu} + A_1$, where λ, μ are rational integers and $|A_1| \leqq 1$. Define

$$\exp(2\pi i A) = \exp(2\pi i \lambda 2^{-\mu})$$

where π and i are the complex numbers usually denoted by these letters. It is easy to see that this is well-defined and that

$$\exp(2\pi i(A+B)) = \exp(2\pi i A)\exp(2\pi i B).$$

Now put

$$\theta_\omega(\alpha) = \exp(2\pi i T(\omega\alpha))$$

for $\omega \in F$, $\alpha \in \mathfrak{o}$, where T denotes the trace from F to Q_2. Then θ_ω is a character defined on $\mathfrak{o}$. Let $\mathfrak{d}$ be the different [1] of the extension F/Q_2. It is easily seen that all $\omega^{-1} \notin \mathfrak{d}$ give trivial characters θ_ω. So we can assume that $\omega^{-1} \in \mathfrak{d}$. Then θ_ω has maximal support $\omega^{-1}\mathfrak{d}^{-1}$. In particular, given any $\mathfrak{p} \subseteq \mathfrak{o}$, there is a θ_ω having $\mathfrak{p}$ as maximal support. Taking $\theta_{\epsilon\omega}$, $|\epsilon| = 1$, instead of θ_ω gives all characters with maximal support $\mathfrak{p}$. Thus every character of $\mathfrak{D}$ can be obtained as a θ_ω for some $\omega^{-1} \in \mathfrak{d}$.

2. The Gauss sums of a lattice. V denotes a non-degenerate n-dimensional vector space with scalar product $x \cdot y$; and L is a lattice on V. For

a fixed $\alpha \in F$, denote by αL the set of vectors αz with $z \in L$; let $\alpha \circ V$ be the vector space V provided with the new metric $x \circ y = \alpha(x \cdot y)$; let $\alpha \circ L$ be the point set L in the space $\alpha \circ V$. It is assumed that $N(L) \subseteq \mathfrak{o}$; i.e., *$L$ will be integral.* These concepts are defined in [8]; we use the same notation here as in the introduction of [8].

For $r \geqq 0$, put $L(r) = \pi^r L$; e.g., $L(0) = L$. Call $x_1, \cdots$ a complete set of representatives of L modulo π^r if

(i) $$x_i - x_j \in L(r) \Longleftrightarrow x_i = x_j,$$

(ii) for each $x \in L$ there is an x_i such that $x - x_i \in L(r)$.

In effect, the x_i are any set of representative vectors of $L/L(r)$. If $y_1, \cdots$ is another complete set of representatives of L modulo π^r, then the y_i can be reordered in such a way that $x_i - y_i \in L(r)$. In particular, these two sets have the same number of elements. By expressing L in a minimal basis we see that any such set of representatives contains 2^{nfr} vectors. Now let $z_1, \cdots$ be a complete set for L modulo π^{q-r}, with $q \geqq r$. Then it is easy to prove that $x_i + \pi^r z_j$ is a complete set for L modulo π^q.

We define the *Gauss sum*

(5) $$\chi(L;\pi^r) = \sum_{x \bmod L(r)} \chi(x^2) \qquad \text{for } \chi \in \mathfrak{D}(r).$$

For short, put

(6) $$\chi(L) = \chi(L;\pi^{m(\chi)}).$$

The definitions are independent of the choice of representatives of $L/L(r)$.

PROPOSITION 1. *If $L = J \oplus K$, then $\chi(L) = \chi(J)\chi(K)$.*

Proof. $$\chi(L) = \Sigma\Sigma\chi((x+y)^2) = \Sigma\Sigma\chi(x^2+y^2) = \Sigma\Sigma\chi(x^2)\chi(y^2)$$
$$= \Sigma\chi(x^2)\Sigma\chi(y^2) = \chi(J)\chi(K),$$

with $x \bmod J(m)$, $y \bmod K(m)$. q. e. d.

PROPOSITION 2. $\chi(L;\pi^r) = 2^{nf(r-m(\chi))}\chi(L)$.

Proof. $$\chi(L;\pi^r) = \Sigma\Sigma\chi((x+\pi^m y)^2) = \Sigma\Sigma\chi(x^2 + 2\pi^m x\cdot y + \pi^{2m}y^2)$$
$$= \Sigma\chi(\pi^{2m}y^2)\Sigma\chi(x^2) = 2^{fn(r-m)}\chi(L)$$

with $x \bmod L(m)$ and $y \bmod L(r-m)$. q. e. d.

For $r \geqq 0$, let the *representation number* $L(h;\pi^r)$ be defined as the number of times $h \in \mathfrak{o}$ is represented modulo π^r by a complete set of representative vectors of $L/L(r)$. This quantity is well-defined.

PROPOSITION 3. *L and K have the same dimension. Then $\chi(L)=\chi(K)$ for all $\chi\in\mathfrak{D}(r)$ if and only if $L(h;\pi^r)=K(h;\pi^r)$ for all $h\in\mathfrak{o}$.*

Proof. Put $\dim L=n=\dim K$. The condition $\chi(L)=\chi(K)$ can be replaced by $\chi(L;\pi^r)=\chi(K;\pi^r)$ in virtue of Proposition 2. In general we have

$$\chi(L;\pi^r)=\sum_{h \bmod \mathfrak{o}(r)} L(h;\pi^r)\chi(h). \tag{7}$$

In particular, this proves the sufficiency. We now prove the necessity.

For a fixed h_0, (7) implies that

$$\chi^{-1}(h_0)\chi(L;\pi^r)=\sum L(h;\pi^r)\chi(h-h_0)+L(h_0;\pi^r),$$

where the summation extends over $h \bmod \mathfrak{o}(r)$, $h\not\equiv h_0 \bmod \mathfrak{o}(r)$. Now summing both sides of this equation over $\chi\in\mathfrak{D}(r)$, we see that (4) implies

$$2^{fr}L(h_0;\pi^r)=\sum_{\chi\in\mathfrak{D}(r)}\chi^{-1}(h_0)\chi(L;\pi^r).$$

This also holds for K. Hence $L(h_0;\pi^r)=K(h_0;\pi^r)$. q. e. d.

Let L be an integral lattice and consider $\pi^k\circ L$ with k a non-negative integer. It is easy to see that

$$\chi(\pi^k\circ L;\pi^r)=\chi_{(k)}(L;\pi^r) \qquad \text{if } \chi\in\mathfrak{D}(r). \tag{8}$$

For all $\chi\in\mathfrak{D}$ we therefore have

$$\begin{aligned} &\chi(\pi^k\circ L)=2^{nfk}\chi_{(k)}(L) && \text{if } 0\leqq k\leqq m(\chi),\\ &\chi(\pi^k\circ L)=2^{nfm(\chi)} && \text{if } m(\chi)\leqq k. \end{aligned} \tag{9}$$

Now consider $\chi\in\mathfrak{D}(r+k)$ with $r\geqq 0$, $k\geqq 0$ fixed; let L and K be n-dimensional integral lattices. Then it follows from (9) that

$$\chi(\pi^k\circ L)=\chi(\pi^k\circ K)\Leftrightarrow\chi_{(k)}(L)=\chi_{(k)}(K). \tag{10}$$

If $\theta\in\mathfrak{D}(r)$, then $\theta=\chi_{(k)}$ for some $\chi\in\mathfrak{D}(r+k)$. So *L and K have the same Gauss sums modulo π^r if and only if $\pi^k\circ L$ and $\pi^k\circ K$ have the same Gauss sums modulo π^{r+k}.*

3. Fundamental invariants of a lattice. Let L and K be two lattices in the same space V. We put

$$L^2=\{x^2\,|\,x\in L\},\qquad x_0\cdot L=\{(x_0\cdot x)\,|\,x\in L\},$$
$$L\cdot K=\{(x\cdot y)\,|\,x\in L \text{ and } y\in K\}.$$

Then $x_0\cdot L$ and $L\cdot K$ are ideals; L^2 is not an ideal, but L^2 generates $N(L)$.

Thus $\alpha \in L^2$ means that L represents α; and $\alpha \in L^2 \bmod \pi^k$ means that $x^2 - \alpha \in \mathfrak{o}(k)$, for some $x \in L$. If $L_1 \oplus L_2 \oplus \cdots \oplus L_t$ is a canonical decomposition of L, we put $L_j \cdot L_j = \mathfrak{o}(s(j))$ and we define $N(L_{(s(j))}) = \mathfrak{o}(u(j))$ where $L_{(s(j))}$ is the invariant substructure given by

$$L_{(s(j))} = \{x \mid x \in L \text{ and } x \cdot L \subseteq \mathfrak{o}(s(j))\}.$$

Define $\mathfrak{o}(v(j)) = \Sigma x^2 \mathfrak{o} + 2\pi^{s(j)}\mathfrak{o}$ where $x \in L_{(s(j))}$ with $\operatorname{ord} x^2$ and $u(j)$ of opposite parity. Then for any $a_j \in L_{(s(j))}{}^2 \bmod \pi^{v(j)}$ with $\operatorname{ord} a_j = u(j)$, it is true that

$$L_{(s(j))}{}^2 \subseteq a_j \mathfrak{o}^2 + \pi^{v(j)}\mathfrak{o}; \tag{11}$$

and $a_j\mathfrak{o}^2 + \pi^{v(j)}\mathfrak{o}$ is an additive group. If L is big, there is equality in (11). We shall call the quantities a_j and $v(j)$ the *fundamental invariants* of L. Then $v(j)$ is unique; while a_j is fixed modulo $\pi^{v(j)}$, but for a factor which is the square of a unit. Every lattice L has a *saturated decomposition* $L = \Sigma L_\lambda$ in which $N(L_j) = \mathfrak{o}(u(j))$, $1 \leqq j \leqq t$. If L is big, it has a *supersaturated decomposition* $L = \Sigma L_\lambda$ in which $L_j{}^2 = a_j\mathfrak{o}^2 + \pi^{v(j)}\mathfrak{o} = L_{(s(j))}{}^2$, for $1 \leqq j \leqq t$. For a detailed discussion, see Sections 1-4 of [8].

4. Unitary lattices. Throughout this section L and K will be n-dimensional π^0-unitary lattices with the same fundamental invariants $v_L = v = v_K$ and $a(L) \cong a \cong a(K) \bmod \pi^v$. We write $u = \operatorname{ord} a$; assume that $\pi^u\mathfrak{o} = a\mathfrak{o} = N(L) = N(K)$. Using results in [8] we shall characterize L in terms of its fundamental invariants and its Gauss sums.

Suppose that n is even and $d(L) \cong (-1)^{\frac{1}{2}n}(1 + [\pi^{u+v}])$. We contend that

$$L \cong \begin{pmatrix} [\pi^u] & 1 \\ 1 & [\pi^v] \end{pmatrix} \oplus \Sigma \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{12}$$

If $n = 2$ this follows by a determinantal consideration. So take $n \geqq 4$ and express L in the semi-canonical decomposition

$$L \cong \begin{pmatrix} [\pi^u] & 1 \\ 1 & \{\pi^v\} \end{pmatrix} \oplus \begin{pmatrix} \pi^v & 1 \\ 1 & \{4\pi^{-v}\} \end{pmatrix} \oplus \Sigma \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{13}$$

given in (12)-(14) of [8]. If u and v have the same parity, again we are through. So take $n \geqq 4$ with $u + v$ odd. Let the basis corresponding to (13) be $\langle x \rangle$. Then we can assume that $\operatorname{ord} x_2{}^2 = v$ since

$$d(L) \cong (-1)^{\frac{1}{2}n}(1 + [\pi^{u+v}]).$$

By successively applying $\operatorname{op}(x_3 \to x_3 + \alpha x_2)$ and $\operatorname{op}(x_3 \to x_3 + \alpha x_1)$ we get the desired result (12).

Now consider the $(n+2)$-dimensional lattices $J \oplus L$, $J \oplus K$ where

$$J \cong \begin{pmatrix} a & 1 \\ 1 & \pi^v \end{pmatrix}.$$

CONTENTION. *If* $\mathfrak{p}$ *is any ideal,* $0 \subseteq \mathfrak{p} \subseteq \mathfrak{o}(v)$, *then*

$$J \oplus L \cong J \oplus K \bmod \mathfrak{p} \Longrightarrow L \cong K \bmod \mathfrak{p}. \tag{14}$$

Proof. Clearly $J \oplus L$, L, K, $J \oplus K$ all have the same fundamental invariants. Then $J \oplus L$ and $J \oplus K$ satisfy the condition of Theorem 1 [8], hence L and K satisfy these conditions, hence $L \cong K \bmod \mathfrak{p}$. q. e. d.

Let H be a π^k-hyperbolic plane, $k \geqq 0$; let χ be any character in $\mathfrak{D}$. Then

$$\chi(H) = \sum_{\alpha \bmod \mathfrak{o}(m)} \sum_{\beta \bmod \mathfrak{o}(m)} \chi(2\pi^k\alpha\beta).$$

But $\chi(2\pi^k\alpha\beta) = \theta(\beta)$ is a character for each fixed α. Hence it follows from (2) that

$$\sum_{\beta \bmod \mathfrak{o}(m)} \chi(2\pi^k\alpha\beta) \geqq 0,$$

with strict inequality occurring for at least one value of α. Hence

$$\chi(H) > 0. \tag{15}$$

THEOREM 1. *L and K are π^0-unitary lattices of the same dimension such that $v_L = v = v_K$ and $a(L) \cong a \cong a(K) \bmod \pi^v$. If $0 \subseteq \mathfrak{p} \subseteq \pi^v\mathfrak{o}$, then $L \cong K \bmod \mathfrak{p}$ if and only if* (i) $d(L) \cong d(K) \bmod a\mathfrak{p}$, (ii) $\chi(L) = \chi(K)$ *for all χ supported by $\mathfrak{p}$.*

Proof. The necessity follows at once from Theorem 1 of [8]. We now prove the sufficiency.

If $|\mathfrak{p}| \geqq |4\pi^{-v}|$, we are through by Theorem 1 of [8]. Hence assume that $|\mathfrak{p}| < |4\pi^{-v}|$ for the rest of the proof. By Proposition 9 of [8] we can assume that L and K are big. Since both L and K now represent a, it suffices to prove that $(a) \oplus L \cong (a) \oplus K \bmod \mathfrak{p}$ when $u = 0$, in virtue of Corollary 9 of [8]. In other words, we can assume that n is even. We have $|a\mathfrak{p}| \leqq |4\pi^{u+1-v}| \leqq |\pi^{u+1+v}|$, and so $d(L) \cong d(K) \bmod \pi^{u+v+1}$. Thus we either have

$$d(L) \cong (-1)^{\frac{1}{2}n}(1 + [\pi^{u+v}]), \qquad d(K) \cong (-1)^{\frac{1}{2}n}(1 + [\pi^{u+v}]) \tag{16}$$

or

$$d(L) \cong (-1)^{\frac{1}{2}n}(1 + \{\pi^{u+v+1}\}), \qquad d(K) \cong (-1)^{\frac{1}{2}n}(1 + \{\pi^{u+v+1}\}). \tag{17}$$

In the latter case adjoin $\begin{pmatrix} a & 1 \\ 1 & \pi^v \end{pmatrix}$ to L and K; then the enlarged lattices

satisfy (16); and it suffices to prove the isometry of the new lattices modulo $\mathfrak{p}$ by (14). In effect we can therefore assume that (16) holds for L and K. By (12) and (15) the problem reduces to the case $n=2$. We assume that this is so.

Since L and K represent the same numbers modulo $\mathfrak{p}$ we can write

$$L \cong \begin{pmatrix} \epsilon\pi^u & 1 \\ 1 & \alpha\pi^v \end{pmatrix}, \qquad K \cong \begin{pmatrix} \epsilon\pi^u + \{\mathfrak{p}\} & 1 \\ 1 & \beta\pi^v \end{pmatrix}$$

with ϵ a unit, α and β integers. Define

$$K^* \cong \begin{pmatrix} \epsilon\pi^u & 1 \\ 1 & \beta\pi^v + \Gamma \end{pmatrix}$$

where Γ is chosen in $\mathfrak{p}$ so that $d(L) \cong d(K^*)$. Then $L \cong K^*$ by Theorem 14.3 of [7], hence $L \cong K \bmod \mathfrak{p}$. q. e. d.

Note. We use the Hasse symbol $S(L)$ in references to [8]. This is defined for the space V by Witt [11] and we put $S(L)=S(V)$. For the general laws obeyed by $S(L)$ see [11]. In addition we have seen in [8] that over fields of our type,

$$(1+4\alpha, \epsilon) = 1, \tag{18}$$

$$(1+4\alpha, \epsilon\pi) = 1 \Leftrightarrow (1+4\alpha)^{\frac{1}{2}} \in F, \tag{19}$$

where $\alpha \in \mathfrak{o}$ and $|\epsilon| = 1$. Also

$$(A,B) = (1+4C, \pi) \quad \text{for some } C \in \mathfrak{o}. \tag{20}$$

And since $\mathfrak{o}/(\pi\mathfrak{o})$ is a finite field it has a quadratic extension, hence there is a unit $E \in \mathfrak{o}$ such that

$$(1+4E)^{\frac{1}{2}} \notin F. \tag{21}$$

Now let L and K be π^0-unitary lattices of the same odd dimension such that $\chi(L) = \chi(K)$ for all $\chi \in \mathfrak{D}(2e+1)$. We shall prove

$$d(L) \cong d(K). \tag{22}$$

Assume that this equation is false for given L and K. Put $L' \cong (1) \oplus (\pi \circ L)$ and $K' \cong (1) \oplus (\pi \circ K)$. Then $\chi(L') = \chi(K')$ for all $\chi \in \mathfrak{D}(2e+2)$. Using the "op" transformations and the operator $\mathfrak{R}$ of [8], or otherwise, it is easily seen that

$$L' \cong [1] \oplus [\pi] \oplus \sum \begin{pmatrix} 0 & \pi \\ \pi & 0 \end{pmatrix}.$$

And similarly with K'. By (15) we can therefore assume that $L' \cong [1] \oplus [\pi]$

and $K' \cong [1] \oplus [\pi]$. By a suitable change of metric on L' and K', we can take $L' \cong 1 \oplus [\pi]$; by a suitable choice of prime π we can take $L' \cong 1 \oplus \pi$. Now $K'(1;4\pi^2) = L'(1;4\pi^2)$. By Hensel's lemma we therefore have

$$L' \cong 1 \oplus \pi, \qquad K' \cong 1 \oplus \pi(1 + G\pi^k) \tag{23}$$

where G is a unit and either (i) $k = 2e$, or (ii) $k < 2e$ with k odd. (i) Since $k = 2e$ with $(1 + G\pi^{2e})^{\frac{1}{2}} \notin F$, it is easily seen that $(1 + G\pi^{2e})\pi \notin (L')^2$, which is impossible since L' and K' represent the same numbers modulo $4\pi^2$. (ii) If $k < 2e$ with k odd, we choose a unit E for which $(1 + 4E)^{\frac{1}{2}} \notin F$. Then there is a unit α such that $\alpha^2 G \equiv E \bmod \pi$. Put $\epsilon = 1 + 4\alpha^2\pi^{-k}$. Then

$$\epsilon \in (L')^2, \qquad \epsilon(1 + 4E) \cong \epsilon + 4\alpha^2 G \in (K')^2.$$

But this implies that $\epsilon(1 + 4E) \in (L')^2$, which is impossible since $L' \cong \epsilon \oplus [\pi]$.

THEOREM 2. *L and K are π^0-unitary lattices of the same dimension. Then $L \cong K$ if and only if $\chi(L) = \chi(K)$ for all $\chi \in \mathfrak{O}(2e + 1)$.*

Proof. Adjoin a π^0-unitary lattice to both L and K so that the enlarged lattices have odd dimension. Then the enlarged lattices have the same determinant by (22). Hence $d(L) \cong d(K)$.

Now $L^2 \equiv K^2 \bmod 4\pi$, hence L and K have the same fundamental invariants, hence $L \cong K \bmod 4\pi$ by Theorem 1, hence $L \cong K$ by Theorem 15.1 of [7]. This proves the sufficiency. The necessity is obvious. q. e. d.

5. The zeros of the Gauss sums. We have seen in (15) that the Gauss sums of a hyperbolic plane never vanish. Thus, for example, it follows that

$$\chi(J) \neq 0 \text{ if } J \cong \begin{pmatrix} 2 & 1 \\ 1 & \{2\} \end{pmatrix}, \tag{24}$$

since $J \oplus J$ is the orthogonal sum of two hyperbolic planes. In this section we solve for χ the equation

$$\chi(L) = 0, \qquad L \text{ fixed and unitary.} \tag{25}$$

LEMMA 1. *L is a π^0-unitary lattice with norm $\mathfrak{o}(u)$. The character χ has maximal support $\mathfrak{o}(m)$ where $0 \leqq m \leqq u$ or $m \geqq 2e - u$. Then $\chi(L) \neq 0$.*

Proof. If this is true for an enlargement of L it is true for L by Proposition 1. Hence assume $\dim L$ even. Since L is now an orthogonal sum of binary lattices of norm $\mathfrak{o}(u)$ it suffices to assume that L is binary, again by Proposition 1. Suppose

$$L \cong \begin{pmatrix} \epsilon\pi^u & 1 \\ 1 & \alpha\pi^v \end{pmatrix} \qquad \alpha \in \mathfrak{o} \tag{26}$$

where v is the second fundamental invariant of L. Using the "op" transformations of [8] we see that

$$\begin{pmatrix} \epsilon\pi^u & 1 \\ 1 & \alpha\pi^v \end{pmatrix} \oplus \begin{pmatrix} -\epsilon\pi^u & 1 \\ 1 & -\alpha\pi^v \end{pmatrix} \cong \begin{pmatrix} [\pi^u] & 1 \\ 1 & 0 \end{pmatrix} \oplus \begin{pmatrix} [\pi^u] & 1 \\ 1 & 0 \end{pmatrix}.$$

Hence the problem reduces to the case where L has the form (26) with $\alpha = 0$.

If $0 \leqq m \leqq u$, then L is a hyperbolic plane modulo $\mathfrak{o}(m)$ and so $\chi(L) \neq 0$ by (15). So assume that $m \geqq 2e - u$. Then

$$\chi(L) = \sum_{\alpha \bmod \mathfrak{o}(m)} \sum_{\beta \bmod \mathfrak{o}(m)} \chi(\epsilon\pi^u\alpha^2)\chi(2\alpha\beta). \tag{27}$$

Now

$$\chi(\epsilon\pi^u\alpha^2) \sum_{\beta \bmod \mathfrak{o}(m)} \chi(2\alpha\beta) = \begin{cases} 0 & \text{if } \alpha \notin \tfrac{1}{2}\mathfrak{o}(m) \\ 2^{fm} & \text{if } \alpha \in \tfrac{1}{2}\mathfrak{o}(m) \end{cases} \tag{28}$$

by (2). Hence $\chi(L) > 0$. q. e. d.

Consider the lattices

$$L \cong \begin{pmatrix} \epsilon\pi^u & 1 \\ 1 & 0 \end{pmatrix}, \quad K \cong \begin{pmatrix} \epsilon\pi^{u+k} & 1 \\ 1 & 0 \end{pmatrix}, \quad |\epsilon| = 1,\ u \geqq 0,\ k \geqq 0, \tag{29}$$

and the character χ with $m(\chi) = m = e + k$. Using (27) and (28) we see that

$$\chi(L) = \Sigma\Sigma\, \chi(\epsilon\pi^u\pi^{2k}\alpha^2)\,\chi(2\pi^k\alpha\beta) = 2^{kf}\,\Sigma\Sigma\, \chi_{(k)}(\epsilon\pi^{u+k}\alpha^2)\chi_{(k)}(2\alpha\gamma)$$

with $\alpha \bmod \mathfrak{o}(m-k)$, $\beta \bmod \mathfrak{o}(m)$, $\gamma \bmod \mathfrak{o}(m-k)$. Hence

$$\chi(L) = 2^{kf}\chi_{(k)}(K). \tag{30}$$

Lemma 2. *Let L have the form (29) with $u < m < 2e - u$. If $u + m$ is even, then* (i) $\chi(L) \neq 0$ *and* (ii) $\chi(L) = 0$ *have solutions with* $m(\chi) = m$. (iii) *If $u + m$ is odd, then $\chi(L) = 0$ for all χ with $m(\chi) = m$.*

Proof. (i) Suppose if possible that $\chi(L) = 0$ for all χ with $m(\chi) = m$. Define

$$J \cong \begin{pmatrix} \epsilon\pi^u & 1 \\ 1 & \pi^{m-1} \end{pmatrix}, \qquad H \cong \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

In virtue of the isometry $J \oplus H \cong L \oplus \cdots$ we have $\chi(J) = 0$. Hence $\chi(J) = \chi(L)$ when $m(\chi) = m$. But $\chi(J) = \chi(L)$ when $\chi \in \mathfrak{O}(m-1) \subseteq \mathfrak{O}(m)$. Hence $L(\pi^{m-1}; \pi^m) = J(\pi^{m-1}; \pi^m)$. This is impossible since $\pi^{m-1} \notin L^2 \bmod \pi^m$.

(ii) Suppose if possible that $\chi(L) \neq 0$ for all χ with $m(\chi) = m$. First take $m \leqq e$. Define

$$J \cong \begin{pmatrix} \epsilon\pi^u + \pi^{m-1} & 1 \\ 1 & 0 \end{pmatrix}, \qquad H \cong \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

In virtue of the relation $L \oplus L \cong L \oplus H$ we have $\chi(L) = \chi(H)$. Since $\chi(L) \neq 0$ for all χ with maximal support $\mathfrak{o}(m)$, it follows that $\chi(\delta \circ L) \neq 0$ for any unit δ. In particular, $\chi(J) \neq 0$ when $m(\chi) = m$. Hence $\chi(J) = \chi(H) = \chi(L)$ when $m(\chi) = m$. For the remaining $\chi \in \mathfrak{O}(m)$, $\chi(J) = \chi(L)$ since $J \cong L \bmod \pi^{m-1}$. Hence $L(\epsilon\pi^u; \pi^m) = J(\epsilon\pi^u; \pi^m)$. This is impossible since $\epsilon\pi^u \notin J^2 \bmod \pi^m$.

Now let $m = e + k$ with $k > 0$. Let K be defined by (29). Using the relations $|2\pi^k| = |\pi^m| > |4\pi^{-u}|$ we see that $|\pi^{u+k}| > |2| > |4\pi^{-(u+k)}|$ and so there is a character χ with $m(\chi) = m$ such that $\chi_{(k)}(K) = 0$. Hence $\chi(L) = 0$ by (30).

(iii) Now let $u + m$ be odd. First suppose that $m \leqq e$. Then

$$\begin{aligned}\chi(L) &= \Sigma\Sigma\, \chi(\epsilon\pi^u\alpha^2)\chi(2\alpha\beta)\\ &= 2^{fm}\, \Sigma\, \chi(\epsilon\pi^u\alpha^2)\\ &= 2^{fm}\, \Sigma\Sigma\Sigma\, \chi(\epsilon\pi^u(a + \pi^{\frac{1}{2}(m-1-u)}b + \pi^{\frac{1}{2}(m+1-u)}c)^2)\\ &= 2^{fm}\, \Sigma\, \chi(\epsilon\pi^u a^2)\, \Sigma\, \chi(\epsilon\pi^{m-1}b^2)\, \Sigma\, \chi(\epsilon\pi^{m+1}c^2)\end{aligned}$$

with

$$\alpha \bmod \mathfrak{o}(m),\ \beta \bmod \mathfrak{o}(m),\ a \bmod \mathfrak{o}(\tfrac{1}{2}(m-1-u)),\ b \bmod \pi,$$
$$c \bmod \mathfrak{o}(\tfrac{1}{2}(m-1+u)).$$

But $\chi_{(m-1)}$ has maximal support $\pi\mathfrak{o}$ and so

$$\Sigma\, \chi(\epsilon\pi^{m-1}b^2) = \Sigma\, \chi_{(m-1)}(\epsilon b^2) = \Sigma\, \chi_{(m-1)}(b) = 0.$$

Hence $\chi(L) = 0$.

Now let $m = e + k$ with $k > 0$. Let K be defined by (29). Then $|\pi^{u+k}| > |2| > |4\pi^{-(u+k)}|$ and so $\chi_{(k)}(K) = 0$. Hence $\chi(L) = 0$ by (30). q. e. d.

Now we can describe the zeros of (25). Let L be a π^0-unitary lattice with fundamental invariants $[\pi^u]$, v, and let $m \geqq 0$ be a fixed rational integer. Then the zeros of $\chi(L)$ with $m(\chi) = m$ are given by

(31)	$m \leqq u$	$\Longrightarrow$	$\chi(L) \neq 0$
(32)	$u < m \leqq v$ with $u + m$ even	$\Longrightarrow$	ambiguous
(33)	$u < m \leqq v$ with $u + m$ odd	$\Longrightarrow$	$\chi(L) = 0$
(34)	$v < m < 2e - v$	$\Longrightarrow$	$\chi(L) = 0$
(35)	$2e - v \leqq m < 2e - u$ with $u + m$ odd	$\Longrightarrow$	$\chi(L) = 0$

$$(36) \qquad 2e - v \leqq m < 2e - u \text{ with } u + m \text{ even} \quad \Longrightarrow \quad \text{ambiguous}$$

$$(37) \qquad 2e - u \leqq m \quad \Longrightarrow \quad \chi(L) \neq 0,$$

where $\chi(L) = 0$ and $\chi(L) \neq 0$ both have solutions in the ambiguous case.

We prove these equations by referring to the two lemmas. The zeros of $\chi(L)$ are unaltered by adjoining hyperbolic planes; hence assume that $\dim L \geqq 4$. Neither the zeros nor the fundamental invariants are changed by adjoining any unit represented by L; hence assume $\dim L$ even. Write

$$L \cong \begin{pmatrix} \epsilon\pi^u & 1 \\ 1 & \alpha\pi^v \end{pmatrix} \oplus \begin{pmatrix} \pi^v & 1 \\ 1 & \{4\pi^{-v}\} \end{pmatrix} \oplus \Sigma \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

As above, we can adjoin $\begin{pmatrix} \epsilon\pi^u & 1 \\ 1 & \alpha\pi^v \end{pmatrix}$ to L and we continue to call the enlarged lattice L. Using a determinantal argument we find

$$L \cong \begin{pmatrix} [\pi^u] & 1 \\ 1 & \{4\pi^{-u}\} \end{pmatrix} \oplus \begin{pmatrix} \pi^v & 1 \\ 1 & \{4\pi^{-v}\} \end{pmatrix} \oplus \Sigma \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Equations (31)-(37) now follow immediately from the two lemmas.

Remark. In the 2-adic case this description of the Gauss sums agrees with the results quoted by Minkowski on p. 57 of [6]. For a discussion of 1-dimensional Gauss sums over number fields see the work of Hecke, especially pp. 218-249 of [4]; compare the results on p. 244 [4] with Lemma 1 of this paper.

6. Unramified theory. We shall see later that it is easy to characterize lattices over local fields in which 2 is a unit and the fact that the Gauss sums never vanish over these fields is responsible for this. The unramified 2-adic theory is almost as easy. To be sure the Gauss sums can vanish, but the zeros take on a very simple form. Thus we apply (31)-(37) to a π^0-unitary lattice L and we find that

$$(38) \qquad \chi(L) = 0 \Longleftrightarrow m(\chi) = 1 \text{ and } N(L) = \mathfrak{o}.$$

Now let us establish the invariants.

THEOREM 3. *Let* $\operatorname{ord} 2 = 1$. *If L and K are totally integral lattices of the same type such that* $\chi(L) = \chi(K)$ *for all* $\chi \in \mathfrak{D}(3 + s(t))$, *then* $L \cong K$.

Proof. First observe that if L and K have canonical decompositions $L = \Sigma L_\lambda$, $K = \Sigma K_\lambda$ in which $L_1 \cong K_1$, then

$$(39) \qquad \chi(L_1{}^\perp) = \chi(K_1{}^\perp) \qquad \text{for } \chi \in \mathfrak{D}(3 + s(t)),$$

where $L = L_1 \oplus L_1{}^{\perp}$, $K = K_1 \oplus K_1{}^{\perp}$. For we can take $s(1) = 0$. If $\chi(L_1) \neq 0$, we are through by Proposition 1. If $\chi(L_1) = 0$, then χ has maximal support $2\mathfrak{o}$ by (38). But $N(L_1{}^{\perp}) = N(K_1{}^{\perp}) \subseteq 2\mathfrak{o}$. Hence $\chi(L_1{}^{\perp}) = \chi(K_1{}^{\perp})$. So (39) holds.

Now proceed to the required result by induction to t. By the cancellation law established in Theorem 5.2 of [7], take $\dim L_1 = \dim K_1$ odd, where $L = \Sigma L_\lambda$, $K = \Sigma K_\lambda$ are canonical decompositions of L and K. Assume that $s(1) = 1$. For $t = 1$ we are through by Theorem 2. Take $t > 1$. It suffices to prove that

$$\bar{L} = L_0 \oplus L \cong K_0 \oplus K = \bar{K} \text{ where } L_0 \cong 1 \cong K_0,$$

again by Theorem 5.2 of [7]. Note that $\chi(\bar{L}) = \chi(\bar{K})$ for all $\chi \in \mathfrak{D}(3 + s(t))$. Using the decompositions (12)-(14) of [8] on the second components of $\bar{L}$ and $\bar{K}$, it is easy to see that

$$L_0 \oplus L_1 \cong [1] \oplus [2] \oplus \Sigma \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}.$$

And similarly with $K_0 \oplus K_1$. Cancelling all superfluous π^1-hyperbolic planes, this allows us to assume that $\bar{L}$ and $\bar{K}$ have 1-dimensional π^1-components. Since $\bar{L}$ and $\bar{K}$ represent the same units modulo 8, we can write

$$\bar{L} = L_0' \oplus L', \qquad \bar{K} = K_0' \oplus K', \qquad L_0' \cong K_0' \cong [1].$$

Then by (39) we have $\chi(L') = \chi(K')$ when $\chi \in \mathfrak{D}(3 + s(t))$. Again we get

$$L' = L_1'' \oplus L'', \qquad K' = K_1'' \oplus K'', \qquad L_1'' \cong K_1'' \cong [2],$$

with $\chi(L'') = \chi(K'')$ when $\chi \in \mathfrak{D}(3 + s(t))$. But L'' and K'' are of the same type. Hence by the inductive assumption $L'' \cong K''$. Hence $\bar{L} \cong \bar{K}$. q. e. d.

7. Characterization by Gauss sums. Let us consider two totally integral lattices L and K; suppose that these lattices are of the same type and that their Gauss sums are equal for all characters in $\mathfrak{D}(2e + s(t) + 1)$. If L and K are unitary this implies that $L \cong K$, and so their Gauss sums are equal for all $\chi \in \mathfrak{D}$. In general, the equivalence class of L cannot be determined by Gauss sums alone, but we shall show that all subsequent Gauss sums can be expressed in terms of the Gauss sums modulo $4\pi^{s(t)+1}$. Thus nothing is gained by considering those $\chi \notin \mathfrak{D}(2e + s(t) + 1)$. Before proving this we shall need some preliminary results.

Consider a character with $m(\chi) \geqq 2e+2$. Contention:

(40) $$\sum_{\alpha \bmod \mathfrak{o}(m)} \chi(\alpha^2) = 2^f \sum_{\alpha \bmod \mathfrak{o}(m-2)} \chi_{(2)}(\alpha^2).$$

In virtue of Proposition 2 it suffices to prove

(41) $$2^f \sum_{\alpha \bmod \mathfrak{o}(m)} \chi(\alpha^2) = \sum_{\alpha \bmod \mathfrak{o}(m)} \chi(\pi^2\alpha^2),$$

and this amounts to proving

(42) $$\chi(K)/\chi(L) = 2^f$$

where $L \cong (-1) \oplus (1)$ and $K \cong (-1) \oplus (\pi^2)$. Now

$$L \cong \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix}, \qquad K \cong \begin{pmatrix} -1 & \pi \\ \pi & 0 \end{pmatrix}.$$

As in (26)-(28) this leads to $\chi(L) = 2^{mf+ef}$. Similarly we can prove that $\chi(K) = 2^{mf+ef+f}$. Hence (42) holds and the contention is proved.

PROPOSITION 4. *L is a totally integral lattice; χ has maximal support $\mathfrak{o}(m) \subseteq 4\mathfrak{o}(s(t)+2)$. Then $\chi(L) = 2^{nf}\chi_{(2)}(L)$.*

Proof. Write L in the canonical decomposition $L = L_1 \oplus \cdots \oplus L_t$. Define the t-dimensional lattice $I = \Sigma I_\lambda$ with $I_\lambda \cong \pi^{s(\lambda)}$. Then $\chi(I) = 2^{tf}\chi_{(2)}(I)$ in virtue of (40) and Proposition 2. Since $I \oplus L$ has an orthogonal basis, it again follows that $\chi(I \oplus L) = 2^{(t+n)f}\chi_{(2)}(I \oplus L)$. Since $\chi(I) \neq 0$, we have $\chi(L) = 2^{nf}\chi_{(2)}(L)$. q. e. d.

COROLLARY 1. *If K has the same type as L and $\chi(L) = \chi(K)$ for all $\chi \in \mathfrak{D}(2e+s(t)+1)$, then $\chi(L) = \chi(K)$ for all $\chi \in \mathfrak{D}$. And $L^2 = K^2$.*

Proof. The first part is an immediate consequence of the proposition. Now take $\alpha \in L^2$. Then $\alpha \in K^2 \bmod 4\pi\alpha$ since $L(\alpha; 4\pi\alpha) = K(\alpha; 4\pi\alpha)$. Hence $\alpha \in K^2$ by Hensel's lemma. q. e. d.

Let us consider two lattices L and K which are of the same type and have the same fundamental invariants a_λ, $v(\lambda)$ for $1 \leqq \lambda \leqq t$. Let J be defined as

$$J \cong \begin{pmatrix} a_1 & \pi^{s(1)} \\ \pi^{s(1)} & \pi^{v(1)} \end{pmatrix}.$$

Then

(43) $$J \oplus L \cong J \oplus K \Longrightarrow L \cong K.$$

This can be verified quite easily by applying both the necessity and sufficiency of Theorem 2 of [8]. The proof is left to the reader.

Consider two lattices $I = I_2 \oplus \cdots \oplus I_t$ and $J = J_2 \oplus \cdots \oplus J_t$ which are of the same type and have the same fundamental invariants a_λ, $v(\lambda)$ for $2 \leqq \lambda \leqq t$; also let $s(2) = 1$. We contend that

$$(44) \qquad 0 \leqq m = m(\chi) < 2e + 2 - v(2) \Longrightarrow \chi(I) = \chi(J).$$

In order to prove this we assume that the given decompositions are big and supersaturated; there is no loss of generality in doing this. Clearly $I_\lambda \cong J_\lambda \bmod \pi^{v(\lambda)}$, and so $I_\lambda \cong J_\lambda \bmod \pi^{v(2)}$ for $2 \leqq \lambda \leqq t$. Hence $\chi(I) = \chi(J)$ when $\mathfrak{o}(m) \supseteq \mathfrak{o}(v(2))$. Now investigate the remaining m, i.e.

$$(45) \qquad |4/\pi^{v(2)-1}| < |\pi^{m-1}| < |\pi^{v(2)-1}|.$$

Then $\chi(I_2) = 2^{n(2)f}\chi_{(1)}(\pi^{-1} \circ I_2)$ in virtue of (9). But $\pi^{-1} \circ I_2$ has fundamental invariants $[\pi^{u(2)-1}]$, $v(2) - 1$; while $\chi_{(1)}$ has maximal support $\mathfrak{o}(m-1)$; hence by (34) and (45) we have $\chi(I_2) = 0$; hence $\chi(I) = 0$ by Proposition 1. By symmetry, $\chi(J) = 0$. This proves the contention.

We shall also need the following. Let I and J be defined by

$$I \cong \begin{pmatrix} a & 1 \\ 1 & 0 \end{pmatrix}, \qquad J \cong \begin{pmatrix} a\pi^2 & \pi \\ \pi & 0 \end{pmatrix}, \qquad a \in \mathfrak{o}.$$

Then

$$(46) \qquad \chi(I) = 0 \Longrightarrow \chi(J) = 0 \text{ if } m(\chi) > e.$$

For let us consider $\alpha \bmod \mathfrak{o}(m-1)$, $\beta \bmod \pi$, $\gamma \bmod \mathfrak{o}(m)$. Then

$$\begin{aligned} \chi(J) &= \Sigma\Sigma\Sigma\, \chi(a\pi^2(\alpha + \pi^{m-1}\beta)^2)\chi(2\pi(\alpha + \pi^{m-1}\beta)\gamma) \\ &= 2^f\, \Sigma\Sigma\, \chi(a\pi^2\alpha^2)\chi(2\pi\alpha\gamma) \\ &= 2^f\, \Sigma\Sigma\Sigma\, \chi(a(\beta + \pi\alpha)^2)\chi(2(\beta + \pi\alpha)\gamma) \end{aligned}$$

in virtue of the relation

$$\chi(a(\beta + \pi\alpha)^2) \sum_{\gamma \bmod \mathfrak{o}(m)} \chi(2(\beta + \pi\alpha)\gamma) = 0 \text{ if } |\beta| = 1,$$

which is an immediate consequence of (2). Hence $\chi(J) = 2^f\chi(I) = 0$.

In particular this proves that if

$$G \cong \begin{pmatrix} a & 1 \\ 1 & \delta\pi^v \end{pmatrix} \text{ with } v = v_G(1),\ |\delta| = 1,$$

then for $m(\chi) \geqq 2e - v$ and $m(\chi) > e$,

$$(47) \qquad \chi(G) = 0 \Longrightarrow \chi(J) = 0.$$

For it can easily be verified, for instance by Theorem 2 of [8], that $G \oplus (-1 \circ G) \cong I \oplus \begin{pmatrix} \pi^v & 1 \\ 1 & 0 \end{pmatrix}$. Hence by (37) and (46) we get $\chi(J) = 0$.

Two definitions. The ideal $\mathfrak{f}_j$, $1 \leqq j \leqq t-1$, is defined as in Section 5 of [8]. The ideal $\mathfrak{g}$ is defined as in equation (2) of [8].

THEOREM 4. *L and K are totally integral lattices of the same type and $L = \Sigma L_\lambda$, $K = \Sigma K_\lambda$ are canonical decompositions. Then L and K are isometric if and only if for all appropriate λ,*

$$\text{(i)} \qquad v_L(\lambda) = v(\lambda) = v_K(\lambda), \qquad a_\lambda(L) \cong a_\lambda(K) \bmod \pi^{v(\lambda)},$$

$$\text{(ii)} \qquad d(L_1 \oplus \cdots \oplus L_\lambda)/d(K_1 \oplus \cdots \oplus K_\lambda) \cong 1 \bmod \mathfrak{f}_\lambda,$$

$$\text{(iii)} \qquad \chi(L) = \chi(K) \text{ for all } \chi \in \mathfrak{D}(2e + s(t) + 1).$$

Proof. The necessity follows from Theorem 2 of [8]. We shall now establish the sufficiency.

Take $a_\lambda(L) = a_\lambda = a_\lambda(K)$. We proceed by induction on the length $s(t) - s(1)$ of L. We can assume that $s(1) = 0$. The induction can be made to start since the theorem has already been proved for lattices of zero length in Theorem 2. By adjoining suitable hyperbolic planes to L and K, we can suppose that L and K are big. If a_1 is a unit, the enlarged lattices $(a_1) \oplus L$ and $(a_1) \oplus K$ satisfy the conditions of the theorem and it suffices to prove the isometry of the enlarged lattices in virtue of Corollary 3 of [8]. In other words we can assume that L_1 and K_1 are of even dimension.

Suppose that $u(1) = e$. If $u(2) = e$, we can find new decompositions $L = \Sigma L_\lambda$, $K = \Sigma K_\lambda$ in which L_1 and K_1 are orthogonal sums of hyperbolic planes. The complements of L_1, K_1 in L, K then satisfy the conditions of the theorem and so we are through. If $u(2) \geqq e + 1$, then $\mathfrak{f}_1 \subseteq 4\pi\mathfrak{o}$, hence $d(L_1) \cong d(K_1)$ in the original decompositions, hence $L_1 \cong K_1$. Induction again gets us through. We shall therefore assume that $u(1) < e$.

The next step in the reduction of the problem is to show that we can take $s(2) = 1$. Suppose $s(2) > 1$. Define N_* as the 12-dimensional lattice consisting of the orthogonal sum of six π^1-hyperbolic planes. Define L^* and K^* by the canonical decompositions

$$L^* = L_1 \oplus N_* \oplus L_2 \oplus \cdots \oplus L_t, \qquad K^* = K_1 \oplus N_* \oplus K_2 \oplus \cdots \oplus K_t.$$

It suffices to prove that $L^* \cong K^*$. To this end we show that L^* and K^* which are still of length $s(t) - s(1)$ also satisfy (i)-(iii). It follows immediately from their definition that a_λ and $v(\lambda)$ have the same values for L^*, K^* as for L, K when $\lambda = 1, 2, \cdots, t$. If a_* and $v(*)$ describe $L_{(1)}{}^*$ they also describe $K_{(1)}{}^*$. We know that

$$|a_1\pi^2| \leqq |a_*| \leqq |a_1|.$$

It is now clear that (i) and (iii) hold for L^*, K^*. (ii) is true when $\lambda = 2, 3, \cdots, t-1$. Let us write $\mathfrak{f}_1^*, \mathfrak{f}_*, \mathfrak{f}_2, \cdots, \mathfrak{f}_{t-1}$ for the $\mathfrak{f}$'s of L^*, K^*. Then it is enough to prove

$$(48) \qquad \mathfrak{f}_1^* \supseteq \mathfrak{f}_1, \qquad \mathfrak{f}_* \supseteq \mathfrak{f}_1.$$

First consider $u(2) = u(1)$ or $u(2) = u(1) + 1$. Using (23)-(24) of [8] it is easily seen that

$$\mathfrak{f}_1 = \pi^{u(1)+v(1)}\mathfrak{o}, \qquad \mathfrak{f}_1^* = \pi^{u(1)+v(1)}\mathfrak{o}, \qquad \pi^2\mathfrak{f}_* = \pi^{u(2)+v(*)}\mathfrak{o}.$$

If $u(2) = u(1)$ we get (48) from the inequality $v(*) \leqq v(1) + 2$; if $u(2) = u(1) + 1$ we get (48) from $v(*) = u(1) + 2$, $u(2) = v(1)$. Now consider $u(2) \geqq u(1) + 2$; then $u(*) = u(1) + 2$, $v(*) \leqq v(1) + 2$, and $\mathfrak{f}_1^* = \pi^{u(1)+v(*)}\mathfrak{o}$. But

$$\mathfrak{f}_1 = [\pi^{u(1)+u(2)}\mathfrak{g}(a_1a_2), \pi^{u(1)+v(2)}, \pi^{u(2)+v(1)}, 2\pi^{\frac{1}{2}(u(1)+u(2))}]$$

$$\pi^2\mathfrak{f}_* = [\pi^{u(*)+u(2)}\mathfrak{g}(a_1a_2), \pi^{u(*)+v(2)}, \pi^{u(2)+v(*)}, 2\pi^{\frac{1}{2}(u(*)+u(2))+1}],$$

the quantities $2\pi^{\frac{1}{2}(u(1)+u(2))}$, $2\pi^{\frac{1}{2}(u(*)+u(2))+1}$ appearing only when $u(1) + u(2)$ is even. So clearly $\mathfrak{f}_* \supseteq \mathfrak{f}_1$; and $\mathfrak{f}_1^* \supseteq \mathfrak{f}_1$ since $|\pi^{u(2)}\mathfrak{g}(a_1a_2)| \leqq |\pi^{v(*)}|$. Thus (48) is generally true. We accordingly assume that the *given* L and K have $s(2) = 1$.

If the decompositions ΣL_λ, ΣK_λ of the given L, K are varied, conditions (i)-(iii) remain valid. Only (ii) needs to be verified and this follows from the necessity of the theorem. In particular, we can assume that ΣL_λ and ΣK_λ are supersaturated.

A final step in the reduction. Call a canonical decomposition $J = \Sigma J_\lambda$ *simple* if all J_λ are big and supersaturated, $2 \leqq \lambda \leqq t$, while J_1 represents $[\pi^{u(1)}]$, $[\pi^{v(1)}]$, and

$$(49) \qquad J_1 \cong \begin{pmatrix} [\pi^{u(1)}] & 1 \\ 1 & \{\pi^{u(1)}\} \end{pmatrix} \oplus \Sigma \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix};$$

we do not insist that J_1 be big; note that

$$(50) \qquad J_1 \cong \begin{pmatrix} [\pi^{u(1)}] & 1 \\ 1 & [\pi^{v(1)}] \end{pmatrix} \oplus \Sigma \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ if } v(1) < e.$$

We shall show that ΣL_λ, ΣK_λ can always be assumed simple. If $u(2) = u(1)$ or $u(2) = u(1) + 1$, using the "op" transformations in conjunction with Proposition 6 of [8], we can find simple decompositions for L and K. Now consider $u(2) = u(1) + 2$. If $v(2) = v(1)$ we proceed as above to simple decompositions for L and K. So let $v(2) > v(1)$. Then $|\mathfrak{f}_1| < |\pi^{u(1)+v(1)}|$. So one of two things can happen. Either

(51) $d(L_1) \cong (-1)^{\frac{1}{2}n(1)}(1+[\pi^{u(1)+v(1)}])$,

$$d(K_1) \cong (-1)^{\frac{1}{2}n(1)}(1+[\pi^{u(1)+v(1)}]),$$

or

(52) $d(L_1) \cong (-1)^{\frac{1}{2}n(1)}(1+\{\pi^{u(1)+v(1)+1}\})$,

$$d(K_1) \cong (-1)^{\frac{1}{2}n(1)}(1+\{\pi^{u(1)+v(1)+1}\}),$$

where $\dim L_1 = \dim K_1 = n(1)$. In the first case ΣL_λ, ΣK_λ are simple by (12). In the second case adjoin $\begin{pmatrix} a_1 & 1 \\ 1 & \pi^{v(1)} \end{pmatrix}$ to both L and K, and we see that the corresponding decompositions of the enlarged lattices satisfy (51) and are therefore simple. Moreover, the enlarged lattices satisfy the conditions of the theorem since their fundamental invariants are unchanged by the adjunction; and it suffices to prove the isometry of the enlarged lattices in virtue of (43). In general therefore, we can assume that ΣL_λ, ΣK_λ are simple. Cancelling off hyperbolic planes allows us to take $\dim L_1 = \dim K_1 = 2$.

So much for the reduction of the problem. To sum up, assume that the sufficiency holds for lattices of length less than $s(t) - s(1)$, and establish it for the simple decompositions $L = \Sigma L_\lambda$, $K = \Sigma K_\lambda$ which satisfy the conditions of the theorem and are such that

(53) $\qquad u(1) < e, \quad s(1) = 0, \quad s(2) = 1, \quad \dim L_1 = \dim K_1 = 2.$

Case 1. $u(2) = u(1)$. By means of the "op" transformations we can write

$$L = L_1' \oplus L_2' \oplus L_3 \oplus \cdots \oplus L_t, \qquad K = K_1' \oplus K_2' \oplus K_3 \oplus \cdots \oplus K_t$$

with L_2', K_2' supersaturated and

(54) $$L_1' \cong \begin{pmatrix} \pi^{v(1)} & 1 \\ 1 & 4\alpha\pi^{-v(1)} \end{pmatrix}, \qquad K_1' \cong \begin{pmatrix} \pi^{v(1)} & 1 \\ 1 & 4\beta\pi^{-v(1)} \end{pmatrix}.$$

Put $L' = L_2' \oplus L_3 \oplus \cdots \oplus L_t$ and similarly define K'. If $u(1) + v(1)$ is even, then $v(1) = e$ and we can assume that $\alpha = \beta = 0$ in (54). Then L', K' satisfy (i)-(iii) and so $L' \cong K'$. From now on take $u(1) + v(1)$ odd. If $v(1) < e$, then $|\pi^{v(2)}| \geqq |\pi^{v(1)+2}| \geqq |4\pi^{-v(1)}|$, and we can again assume that $\alpha = \beta = 0$. Then (i)-(ii) are immediate for L', K'. We must show that $\chi(L') = \chi(K')$, for all $\chi \in \mathfrak{D}(2e + s(t) + 1)$. By (44) we can assume that $m(\chi) \geqq 2e + 2 - v(2)$; then $m(\chi) \geqq 2e - v(1)$ and so $\chi(L_1') \neq 0$; thus $\chi(L') = \chi(K')$ by division, in virtue of Proposition 1. Hence $L' \cong K'$ by the inductive assumption.

To conclude this case we must consider $u(1) + v(1)$ odd, $v(1) = e$.

If $v(1)=v(2)=e$ the result is trivial since we can take $\alpha=\beta=0$. So let $v(2)=v(1)+1=e+1$. If $(1+4\alpha)\cong(1+4\beta)$ we are through. So suppose that $(1+4\alpha)/(1+4\beta)$ is a non-square. We shall prove this impossible. Define $I\cong 1\cong J$. Then $I\oplus(\pi^e\circ L)$ and $J\oplus(\pi^e\circ K)$ have the same Gauss sums modulo 4π. But

$$I\oplus(\pi^e\circ L)\cong(1+4\alpha)\oplus\begin{pmatrix}0 & 2\\ 2 & 0\end{pmatrix}\oplus(\pi^e\circ L').$$

And similarly with K. Hence $(1+4\alpha)\oplus(\pi^e\circ L')$ represents $1+4\beta$. But $(\pi^e\circ L')$ represents no number of even ordinal $\geqq 2e$. So $(1+4\alpha)\oplus(\pi^e\circ L')$ cannot represent $1+4\beta$. This is a contradiction. Thus the first case is proved.

Case 2. $u(2)=u(1)+1$. We have $u(1)<e$; thus $v(2)=u(1)+2$. Using the "op" transformations and Proposition 6 of [8] we can write

$$L=L_1'\oplus L_2'\oplus L_3\oplus\cdots\oplus L_t,\qquad K=K_1'\oplus K_2'\oplus K_3\oplus\cdots\oplus K_t,$$

where L_2', K_2' are supersaturated and

$$L_1'\cong\begin{pmatrix}\pi^{u(1)} & 1\\ 1 & 0\end{pmatrix},\qquad K_1'\cong\begin{pmatrix}\pi^{u(1)} & 1\\ 1 & 0\end{pmatrix}.$$

Let L', K' denote the complements of L_1', K_1' in L, K respectively. Clearly L', K' continue to satisfy (i)-(ii); we must establish (iii). In virtue of (44) we need only consider $m(\chi)\geqq 2e-u(1)$; thus $\chi(L_1')\neq 0$; thus $\chi(L')=\chi(K')$ by division. By the inductive assumption we again get $L'\cong K'$. This proves the second case.

Case 3. $u(2)=u(1)+2$. We can assume that $a_1\in L^2$; then $a_1\in K^2$ by Corollary 1. We can assume that $u(1)<e$, $a_2=a_1\pi^2$, $v(2)\geqq v(1)\geqq v(2)-2$. Since ΣL_λ and ΣK_λ are simple we have

(55) $\qquad d(L_1)\cong-(1+[\pi^{u(1)+v(1)}]),\qquad d(K_1)\cong-(1+[\pi^{u(1)+v(1)}])$

when $u(1)+v(1)$ is odd. Now L represents a_1, and L_2 represents all numbers represented by $L_2\oplus\cdots\oplus L_t$. So it is easily seen that we can write

$$L_1\oplus L_2=\bar{L}_1\oplus\bar{L}_2\text{ with }\bar{L}_1\cong\begin{pmatrix}a_1 & 1\\ 1 & \{\pi^{v(1)}\}\end{pmatrix}$$

where $\bar{L}_2$ is still a supersaturated component. Similarly define $\bar{K}_1\oplus\bar{K}_2$

$= K_1 \oplus K_2$. If $v(1) = v(2)$ or if $u(1) + v(1)$ is even, we easily see that we can take

$$(56) \qquad \bar{L}_1 \cong \begin{pmatrix} a_1 & 1 \\ 1 & [\pi^{v(1)}] \end{pmatrix}, \qquad \bar{K}_1 \cong \begin{pmatrix} a_1 & 1 \\ 1 & [\pi^{v(1)}] \end{pmatrix}.$$

While if $v(2) > v(1)$ with $u(1) + v(1)$ odd, we have $|\,\mathfrak{f}_1\,| < |\,\pi^{u(1)+v(1)}\,|$, thus $\bar{L}_1$, $\bar{K}_1$ satisfy (55), thus (56) is true for suitably chosen $\bar{L}_1$, $\bar{K}_1$. Write

$$(57) \qquad \bar{L}_1 \cong \begin{pmatrix} a_1 & 1 \\ 1 & \delta\pi^{v(1)} \end{pmatrix} \text{ with } |\,\delta\,| = 1.$$

Since $d(\bar{L}_1) \cong d(\bar{K}_1) \bmod \pi^{u(1)+v(2)}$, using Theorem 14.3 of [7] we can write

$$\bar{K}_1 \cong \begin{pmatrix} a_1 & 1 \\ 1 & \delta\pi^{v(1)} + \{\pi^{v(2)}\} \end{pmatrix}.$$

Using Proposition 6 of [8], we can find a decomposition $K_1' \oplus K_2' = \bar{K}_1 \oplus \bar{K}_2$ in which K_2' is still supersaturated, and $K_1' \cong \bar{L}_1'$. In other words, we have two simple decompositions $L = \Sigma L_\lambda'$, $K = \Sigma K_\lambda'$ in which $L_1' \cong K_1'$ have the form of (57). If L', K' are the complements of L_1', K_1' in L, K respectively, we must show that L', K' satisfy the conditions of the theorem. Once this is verified the isometry $L \cong K$ follows by the inductive assumption. (i) and (ii) are easily checked for L', K'; we must establish (iii) for these lattices.

If $m(\chi) < 2e + 2 - v(2)$, we have $\chi(L') = \chi(K')$ by (44). Now let $m(\chi) \geqq 2e + 2 - v(2)$. Then $m(\chi) \geqq 2e - v(1)$ and $m(\chi) > e$, since $2 + v(1) \geqq v(2)$ and $v(2) \leqq e + 1$. If $\chi(L_1') = \chi(K_1') = 0$, we see that L_2' and K_2' each have a component of the form

$$\begin{pmatrix} a_1\pi^2 & \pi \\ \pi & 0 \end{pmatrix}$$

and so $\chi(L') = 0 = \chi(K')$ in virtue of (47). While if $\chi(L_1') = \chi(K_1') \neq 0$, then $\chi(L') = \chi(K')$ by division. Hence L' and K' have the same Gauss sums. This proves the theorem.

In general it is impossible to determine the integral equivalence of lattices by Gauss sums alone, and this point will be illustrated in the next section. All that can be said about lattices with the same Gauss sums is that they are fractionally equivalent; obviously the converse does not hold.

THEOREM 5. *L and K are totally integral lattices of the same type and $\chi(L) = \chi(K)$ for all $\chi \in \mathfrak{D}(2e + s(t) + 1)$. Then L and K are fractionally equivalent.*

Proof. In virtue of Witt's cancellation theorem [11] we can adjoin as many 1-dimensional lattices to L and K as we please. So assume that we

have canonical decompositions $L=\Sigma L_\lambda$, $K=\Sigma K_\lambda$ in which each component is proper with $s(\lambda+1)=s(\lambda)+1$. The fundamental invariants can then be chosen as

$$a_\lambda=\pi^{s(\lambda)},\qquad v(\lambda)=s(\lambda)+1,\qquad \mathfrak{f}_\lambda=\pi\mathfrak{o},$$

for both L and K. Hence $L\cong K$ by Theorem 4. q. e. d.

8. Remarks. By way of example we shall make some rough estimates of the Gauss sums. χ will be a fixed character with maximal support $\mathfrak{o}(m)$. We put

$$h=2^{2mf}\text{ if } m\leqq e,\qquad h=2^{mf+ef}\text{ if } m\geqq e.$$

Refining the method used to derive (15) it is easy to see that $\chi(H)=h$ for the unit hyperbolic plane H. Now consider the π^0-unitary lattices

$$J_0\cong\begin{pmatrix}a&1\\1&0\end{pmatrix},\qquad J\cong\begin{pmatrix}a&1\\1&b\end{pmatrix}.$$

Then $\chi(J_0)=h$ whenever $\chi(J_0)\neq 0$, since $J_0\otimes J_0\cong J_0\otimes H$. Using this result in conjunction with the isometry

$$J\oplus J\oplus J\oplus J\cong\begin{pmatrix}\alpha&1\\1&0\end{pmatrix}\oplus\begin{pmatrix}\beta&1\\1&0\end{pmatrix}\oplus\begin{pmatrix}\alpha&1\\1&0\end{pmatrix}\oplus\begin{pmatrix}\beta&1\\1&0\end{pmatrix},$$

we get $\chi(J)=\omega_4 h$ whenever $\chi(J)\neq 0$, where ω_4 denotes some complex 4-th root of unity. So if I is a 1-dimensional π^0-unitary lattice, we have $\chi(I)=\omega_8 h^{\frac{1}{2}}$, where ω_8 is an 8-th root of unity. Thus for the n-dimensional π^0-unitary lattice L we have

$$\chi(L)=0\text{ or }\chi(L)=\omega h^{\frac{1}{2}n}\tag{58}$$

where ω is one of the complex 8-th roots of unity.

Now three examples to show the part played by the various assumptions in Theorems 3 and 4. First consider the 2-adic forms

$$L\cong 1\oplus 2\oplus 2\oplus -2\oplus 4,\qquad K\cong 1\oplus 1\oplus 2\oplus -4\oplus 4.$$

Clearly these two forms are not of the same type. Nevertheless they have the same Gauss sums. To see this consider the character χ with maximal support $\mathfrak{p}\subseteq\pi\mathfrak{o}$. For $\mathfrak{p}\supseteq 8\mathfrak{o}$ we have $\chi(L)=\chi(K)=0$. For $\mathfrak{p}\supseteq 16\mathfrak{o}$ it is easily verified that the two lattices

$$2\oplus -2\cong\begin{pmatrix}2&2\\2&0\end{pmatrix}\quad\text{and}\quad 1\oplus -4\cong\begin{pmatrix}1&2\\2&0\end{pmatrix}$$

have the same Gauss sums when computed for the given χ. So the assertion is proved.

Now consider any ramified extension of the 2-adic numbers, with ramification index $e \geqq 6$. Define

$$I \cong \begin{pmatrix} \pi & 1 \\ 1 & 0 \end{pmatrix} \oplus (1), \qquad \cong \begin{pmatrix} \pi & 1 \\ 1 & 0 \end{pmatrix} \oplus (1+\pi), \qquad K \cong \begin{pmatrix} \pi^3 & 1 \\ 1 & 0 \end{pmatrix} \oplus (1).$$

Then the lattices

$$M = I \oplus (\pi^2 \circ J) \quad \text{and} \quad N = J \oplus (\pi^2 \circ I)$$

do not satisfy condition (ii) of Theorem 4. We prove that their Gauss sums are equal. For $\mathfrak{p} \supseteq 4\pi\mathfrak{o}$ we have $\chi(M) = \chi(N) = 0$. For $\mathfrak{p} \subseteq 4\pi^2\mathfrak{o}$ we have $\chi(M) = \chi(N)$ in virtue of the isometry

$$(1) \oplus (-\pi) \oplus \pi^2(1+\pi) \cong (1+\pi) \oplus (-\pi) \oplus (\pi^2).$$

There is another example which shows that the equality of the Gauss sums does not necessarily imply the equality of the fundamental invariants. Consider

$$P = I \oplus (\pi^2 \circ I) \oplus (\pi^4 \circ I), \qquad Q = I \oplus (\pi^2 \circ K) \oplus (\pi^4 \circ I).$$

Then $v_P(2) \neq v_Q(2)$. But $\chi(P) = \chi(Q) = 0$ when $\mathfrak{p} \supseteq 4\pi^3\mathfrak{o}$. And if $\mathfrak{p} \subseteq 4\pi^4\mathfrak{o}$ we have $\chi(P) = \chi(Q)$ in virtue of the relation

$$\begin{pmatrix} \pi^3 & 1 \\ 1 & 0 \end{pmatrix} \oplus \begin{pmatrix} -\pi & 1 \\ 1 & 0 \end{pmatrix} \cong \begin{pmatrix} \pi & 1 \\ 1 & 0 \end{pmatrix} \oplus \begin{pmatrix} -\pi & 1 \\ 1 & 0 \end{pmatrix}.$$

9. At the odd primes. The discussion of the theory over fields in which 2 is a unit is almost trivial. Thus, consider a local field F whose residue class field is a finite field containing p^f elements, $p > 2$. Again we can define the character group $\mathfrak{O}$. For any integral lattice, define the Gauss sums $\chi(L; \pi^r)$ and $\chi(L)$ as in Section 2. Propositions 1-3 are obviously valid. If I is a 1-dimensional π^0-unitary lattice, $-I \oplus I$ is a hyperbolic plane H; it is easily seen that $\chi(H) > 0$; hence $\chi(I) \neq 0$.

THEOREM 6. *L and K are n-dimensional integral lattices over a ring $\mathfrak{o}$ in which 2 is a unit. Then $K \cong L$ if and only if $\chi(K) = \chi(L)$ for all $\chi \in \mathfrak{O}(s_L(t) + 1)$.*

Proof. Induction to n. Write $L = I \oplus I^{\perp}$ with $N(L) = N(I)$ and $I \cong a_1$. Then $a_1 \in K^2 \bmod \pi a_1$. Hence $a_1 \in K^2$ by Hensel's lemma. But $N(K) = N(L) = a_1\mathfrak{o}$ since L and K represent the same numbers modulo

$\mathfrak{o}(s_L(t)+1)$. Hence there is a $J \cong I$ such that $K = J \oplus J^{\perp}$. But $\chi(J^{\perp}) = \chi(I^{\perp})$ for all $\chi \in \mathfrak{D}(s_L(t)+1)$, by division. Hence $J^{\perp} \cong I^{\perp}$. Hence $K \cong L$. q. e. d.

PRINCETON UNIVERSITY.

REFERENCES.

[1] E. Artin, *Algebraic numbers and algebraic functions I*, Princeton University and New York University, 1950-1951.

[2] W. H. Durfee, "Congruence of quadratic forms over valuation rings," *Duke Mathematical Journal*, vol. 11 (1944), pp. 687-697.

[3] E. Hecke, "Reziprozitätsgesetz und Gauss'sche Summen in quadratischen Zahlkörpern," *Gött. Nachr., Math.-phys. Kl.*, pp. 265-278, 1919.

[4] ———, *Vorlesungen über die Theorie der algebraischen Zahlen*, Leipzig, 1923.

[5] B. W. Jones, "A canonical quadratic form for the ring of 2-adic integers," *Duke Mathematical Journal*, vol. 11 (1944), pp. 715-727.

[6] H. Minkowski, "Théorie des formes quadratiques," *Gesammelte Abhandlungen*, Leipzig and Berlin, 1911.

[7] O. T. O'Meara, "Quadratic forms over local fields," *American Journal of Mathematics*, vol. 77 (1955), pp. 87-116.

[8] ———, "Integral equivalence of quadratic forms in ramified local fields," *American Journal of Mathematics*, vol. 79 (1957), pp. 157-186.

[9] G. Pall, "The arithmetical invariants of quadratic forms," *Bulletin of the American Mathematical Society*, vol. 51 (1945), pp. 185-197.

[10] B. L. van der Waerden, *Modern Algebra*, vol. 2, New York, 1950.

[11] E. Witt, "Theorie der quadratischen Formen in beliebigen Körpern," *Journal für die reine und angewandte Mathematik*, vol. 176 (1937), pp. 31-44.

ON THE EXISTENCE OF AN ABSOLUTE CONSTANT CONCERNING "FLAT" OSCILLATORS.*

By Aurel Wintner.

For large positive t, let $\omega(t)$ be a positive continuous function corresponding to which the differential equation (*) $x'' + \omega^2(t)x = 0$ has the following property: $x'(t) \to 0$, as $t \to \infty$, holds for the derivative of every solution $x(t)$ of (*). Then (*) will be called *flat* (all of its solutions "flatten out"; still, they can be "wobbly" and "large," as shown by the example $\omega(t) = c/t$, mentioned below for $c > \frac{1}{2}$).

Since $x(t) = o(t)$ is necessary for $x'(t) = o(1)$, the "frequency" $\omega(t) > 0$ cannot be "too small" (for large t) if (*) is flat. In fact, (*) will have solutions satisfying $x(t) \sim t$ if the integral of $t\omega^2(t)$ over const. $\leqq t < \infty$ is finite (Bôcher); for instance, if $\omega(t) = O(t^{-1-\epsilon})$ for some $\epsilon > 0$. Crucial proves to be the limiting case $\epsilon = 0$, since the issue then involves the numerical value of the constant absorbed by the O of the assumption $\omega(t) = O(t^{-1})$. Put therefore $[\omega] = \limsup t\omega(t)$, and assume that $[\omega] < \infty$.†

Let α be the greatest universal constant having the property that (*) is non-oscillatory whenever $[\omega] < \alpha$, and let β be the greatest universal constant (if any) having the property that (*) is flat whenever $[\omega] < \beta$. It is well-known that $\alpha = \frac{1}{2}$ (cf., e.g., pp. 231-233 Bieberbach's *Differentialgleichungen* of 1956). If $\omega(t) = c/t$, where $c > \frac{1}{2}$, then the solutions $x(t)$ of (*) are superpositions of the real and imaginary parts of $t^{\frac{1}{2}} \exp(iC \log t)$, where $C = C(c)$ is positive (and tends to 0 as $c \to \infty$, and to ∞ as $c \to \frac{1}{2} + 0$; if $c = \frac{1}{2}$, these solutions $x(t)$ degenerate into $t^{\frac{1}{2}}$ and $t^{\frac{1}{2}} \log t$, so that (*) ceases to be oscillatory).

For this explicity integrable "slow oscillator," $x'(t)$ is $O(t^{-\frac{1}{2}}) = o(1)$,

* Received February 26, 1957.

† If $[\omega] < \infty$, then $x''(t) = o(t^{-2})x(t)$, by (*), hence $x''(t) = o(t^{-1})$ if $x(t) = o(t)$. It follows therefore from a well-known inequality (which goes back to Hadamard; cf. p. 12 of Carleman's *Fonctions quasi analytiques*, Paris, 1926) that if $[\omega] < \infty$, then $x'(t) = o(1)$ whenever $x(t) = o(t)$. Accordingly, if $[\omega] < \infty$, then (*) is flat *if and only if* $x(t) = o(t)$ holds for all solutions of (*).

hence (*) is flat; and every $[\omega] > \frac{1}{2}$ is allowed, since $[\omega] = c$, where $\frac{1}{2} < c < \infty$. Thus, in view of the definitions of α and β, one might expect that either $\beta = \alpha$ or $\beta = \infty$. But a theorem of Boas, Boas and Levinson (Duke Math. Journ., vol. 8 (1942), p. 849) implies that (*) is flat whenever $[\omega] < 1$ (provided that (*) is oscillatory; otherwise they conclude from $[\omega] < 1$ the existence of a finite limit $x'(\infty)$, which need not be 0). Hence $\beta \geqq 1$, while $\alpha = \frac{1}{2}$, and so $\beta \neq \alpha$. Consequently, the expected alternative reduces to $\beta = \infty$.

It turns out, however, that $\beta = \infty$ is false; in other words, that *an oscillator* (*) *need not be flat if* $\omega(t) = O(t^{-1})$. In order to shorten the formulae, only $\beta < \infty$ will be proved; but the approach to be followed (patterned after my note in the Journ. of Appl. Physics, vol. 18 (1947), pp. 941-942, where a question in "high frequency" $\omega(t)$, rather than the finer question of the present "low frequency" $\omega(t)$, is considered) is flexible enough to lead to $\beta < 1 + \epsilon$ for every $\epsilon > 0$ and therefore, since $\beta \geqq 1$, to $\beta = 1$.

Try to satisfy (*) by $x(t) = r\cos(C\log t)$, where $C = \text{const.} > 0$, $1 \leqq t < \infty$ and $r = r(t)$. Direct substitution leads to the following determination of $\omega(t)$ in terms of $r(t)$:

$$\omega^2 = -r''/r + \{2t^{-1}Cr'/r - Ct^{-2}\}\tan + C^2/t^2, \tag{1}$$

where the argument of tan (and, subsequently, of cos and sin) is $C\log t$, that of the cos in $x(t)$. In order to keep $\omega^2(t)$ continuous, the coefficient $\{\ \}$ of tan must be so chosen as to cancel the poles of tan.

Such a choice is $\{\ \} = \phi\cos$, if $\phi = \phi(t)$ is continuous. Choose $\phi(t) = 6t^{-2}\cos$. Then (1) reduces to

$$t^2\omega^2(t) = -t^2r''/r + 6\sin\cos + C^2, \tag{2}$$

and $r = r(t)$ follows from $\{\ \} = 6t^{-2}\cos^2$ by a quadrature:

$$r(t) = \exp\int^t \tfrac{1}{2}s^{-1}(1 + 6C^{-1}\cos^2(C\log s))\,ds. \tag{3}$$

Since the substitution $u = C\log s$ shows that $\int^t s^{-1}\cos^2(C\log s)\,ds \sim \frac{1}{2}C\log t$ as $t \to \infty$, it is seen from (3) that $\log r(t) \sim 2\log t$, hence $r(t) \neq o(t)$. It follows therefore from $x(t) = r(t)\cos$ that $x(t) \neq o(t)$. Hence $x'(t) \neq o(1)$, which means that (*) cannot be flat.

On the other hand, two differentiations of (3) show that

$$t^2r''/r = (\tfrac{1}{2} + 3C^{-1}\cos^2)^2 - (\tfrac{1}{2} + 3C^{-1}\cos^2) - 6\sin\cos. \tag{4}$$